

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:27:33 ; Search time 34.9726 seconds

(Without alignments)
2434.679 Million cell updates/sec

Title: US-09-920-954-1

Sequence: 1 MRKKVFLSVLSAAAILSTV.....EYQATNPVPGQXSLATVN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3030	97.7	639	20	AA17089
2	3022	97.5	640	20	AA17090
3	3021	97.4	640	20	AA17091
4	3017	97.3	639	20	AA17087
5	3007	97.0	640	20	AA17088
6	2723	87.8	641	20	AAW89547
7	2408	77.7	636	20	AAW89548
8	2155	69.5	434	23	AAW50080
9	2155	69.5	434	23	AAW50081
10	2082	67.1	434	23	AAW50085

11	2060.5	66.4	433	23	AAW50086	Bacillus sp alkali
12	1994	64.3	434	23	AAW50090	Bacillus sp KSM-KP
13	1952.5	63.0	433	23	AAW50084	Bacillus sp SD-521
14	1948.5	62.8	433	23	AAW50082	Bacillus sp D6-(FE
15	1941.5	62.6	433	23	AAW50083	Bacillus sp Y-(FER
16	1940.5	62.6	433	13	AAW62774	Alkali-protase Ya
17	1940.5	62.6	433	19	AAW61495	Modified Bacillus
18	1940.5	62.6	433	21	AAW95698	Bacillus sp. Lion
19	1940.5	62.6	433	21	AAW94207	Amino acid sequenc
20	1940.5	62.6	433	21	AAW44619	Bacillus lion y en
21	1644.5	53.0	345	20	AAW62230	Subtilase jpi70 fr
22	1644.5	53.0	345	20	AAW1654	Thermococcus prote
23	451.5	14.6	659	18	AAW24121	MO9856926 Seq ID 1
24	451.5	14.6	659	20	AAW24129	Pyrococcus furiosu
25	408	13.2	654	18	AAW24129	Pyrococcus furiosu
26	408	13.2	654	18	AAW24129	Pyrococcus furiosu
27	398	12.8	659	18	AAW24123	Protease. Synthetic
28	391	12.6	522	18	AAW24122	Hyperthermostable
29	391	12.6	522	18	AAW24122	Hyperthermostable
30	377.5	12.2	545	22	AAW94838	T. yonsei subcll
31	377.5	12.2	545	22	AAW94838	Transglutaminase r
32	358	11.5	1079	22	AAW1180	Fragment of dhp a g
33	340	11.0	734	18	AAW13667	Streptomyces virid
34	340	11.0	734	18	AAW13667	DhpA-mel chimeric
35	340	11.0	734	18	AAW13667	Bacillus sp. Ty145
36	307.5	9.9	903	17	AAW87007	Hyperthermostable
37	299.5	9.7	1398	17	AAW87008	Protease. Pyrococ
38	299.5	9.7	1398	17	AAW87008	Pyrococcus furiosu
39	299.5	9.7	1398	18	AAW24124	MO9856926 Seq ID 6
40	299.5	9.7	1398	20	AAW4839	F. balustinum CP70
41	292	9.4	699	20	AAW08471	Hyperthermostable
42	288.5	9.3	237	17	AAW87009	RP-III residual pr
43	286	9.2	806	13	AAW27481	Peptide sequence.
44	283.5	9.1	188	17	AAW87011	Thermococcus prote
45	283.5	9.1	188	18	AAW24126	

ALIGNMENTS

```
RESULT 1
AA17089
ID AA17089 standard: Protein: 639 AA.
XX
XX AA17089;
AC
XX 21-JUL-1999 (first entry)
DT
XX
XX
DE Bacillus alkaline protease.
XX
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidizing agent.
XX
XX
XX Bacillus sp.
OS
XX
XX
XX WO9918218-A1.
XX
XX 15-APR-1999.
PD
XX
XX 07-OCT-1998; 98WO-JP04528.
XX
XX 07-OCT-1997; 97JP-0274570.
XX
XX (KAOS ) KAO CORP.
XX Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
XX Saeki K, Shikata S, Takaiwa M;
XX
XX WPI: 1999-287736/24.
XX N-PSDB: AAX37277.
XX
XX Alkali protease from Bacillus used in washing powders
XX
```

PS Disclosure; Page 53-58; 71pp; Japanese.
 XX
 CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease.
 XX
 SO Sequence 639 AA;
 Query Match 97.7%; Score 3030; DB 20; Length 639;
 Best Local Similarity 93.4%; Pred. No. 4.2e-238;
 Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 OY 1 MRKKVFLSVLSAAALSTVVALXNPAGARXEDLDKGIQTDTDXGFSKXOTGAAPL 60
 DB 1 MRKKVFLSVLSAAALSTVVALXNPAGARXEDLDKGIQTDTDXGFSKXOTGAAPL 60
 OY 61 LVESENVKLKGLKKLETPANNKLIHQFNGPILEETKQXLEXTGAKLIDYIPYAYI 120
 DB 61 LVESENVKLKGLKKLETPANNKLIHQFNGPILEETKQXLEXTGAKLIDYIPYAYI 120
 OY 121 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 180
 DB 121 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 180
 OY 121 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 180
 DB 121 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 180
 OY 181 IETIAQXXXNDYXITTAPEYKVMNDVARGIVKADVAOSYGLYOGQIVAAADGLDT 240
 DB 181 IETIAQXXXNDYXITTAPEYKVMNDVARGIVKADVAOSYGLYOGQIVAAADGLDT 240
 OY 181 IETIAQXXXNDYXITTAPEYKVMNDVARGIVKADVAOSYGLYOGQIVAAADGLDT 240
 DB 181 IETIAQXXXNDYXITTAPEYKVMNDVARGIVKADVAOSYGLYOGQIVAAADGLDT 240
 OY 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVRO 300
 DB 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVRO 300
 OY 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVRO 300
 DB 241 GRNDSMHEAFRGKITALYALGRTNNAANDTNGHGTAVAGSVLNGXTNKGMAPQANLVRO 300
 OY 301 STMDSSGGLGSLNQLTFLSQAXSAGARIHTNSMGAANYGATTTSSRNVDYVRKNDMT 360
 DB 301 STMDSSGGLGSLNQLTFLSQAXSAGARIHTNSMGAANYGATTTSSRNVDYVRKNDMT 360
 OY 301 STMDSSGGLGSLNQLTFLSQAXSAGARIHTNSMGAANYGATTTSSRNVDYVRKNDMT 360
 DB 301 STMDSSGGLGSLNQLTFLSQAXSAGARIHTNSMGAANYGATTTSSRNVDYVRKNDMT 360
 OY 361 IIFPAENEXPNNGTISAPGAKNAITVGATENLRPSFGSADININVAOPSSRGPTKDR 420
 DB 361 IIFPAENEXPNNGTISAPGAKNAITVGATENLRPSFGSADININVAOPSSRGPTKDR 420
 OY 361 IIFPAENEXPNNGTISAPGAKNAITVGATENLRPSFGSADININVAOPSSRGPTKDR 420
 DB 361 IIFPAENEXPNNGTISAPGAKNAITVGATENLRPSFGSADININVAOPSSRGPTKDR 420
 OY 421 IRPDVAPGTXLISARSSLAPDSEFWANHDSKYAYMGTSMATPIVAGVNAQUREHFVN 480
 DB 421 IRPDVAPGTXLISARSSLAPDSEFWANHDSKYAYMGTSMATPIVAGVNAQUREHFVN 480
 OY 421 IRPDVAPGTXLISARSSLAPDSEFWANHDSKYAYMGTSMATPIVAGVNAQUREHFVN 480
 DB 421 IRPDVAPGTXLISARSSLAPDSEFWANHDSKYAYMGTSMATPIVAGVNAQUREHFVN 480
 OY 481 RGTTPPSLLKALLAGAADXGLGYNGNOGWRVTLKSLNVAAYNESSALSTOKATY 540
 DB 481 RGTTPPSLLKALLAGAADXGLGYNGNOGWRVTLKSLNVAAYNESSALSTOKATY 540
 OY 541 XPFATGKRLKLSIWSADPASTTASVTLVNDLDTITPANGXHYXNGNFXKXXXNMNG 600
 DB 541 XPFATGKRLKLSIWSADPASTTASVTLVNDLDTITPANGXHYXNGNFXKXXXNMNG 600
 OY 601 RNNVENVFINSPOSGTYTIEVOAYNPVGPQNFSLAIYN 639
 DB 601 RNNVENVFINSPOSGTYTIEVOAYNPVGPQNFSLAIYN 639
 RESULT 2
 AA17090
 ID AA17090 standard; Protein; 640 AA.
 XX AA17090;
 XX

DT 21-JUL-1999 (first entry)
 XX
 DE Bacillus alkaline protease.
 XX
 KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidising agent.
 OS Bacillus sp.
 XX
 PN MO9918218-AL.
 XX
 PD 15-APR-1999.
 XX
 PF 07-OCT-1998; 98WO-JP04528.
 XX
 PR 07-OCT-1997; 97JP-0274570.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
 PI Saeiki K, Shikata S, Takaiwa M;
 XX
 DR WPI: 1999-287336/24.
 XX
 DR N-PSDB: AAX37278.
 XX
 PT Alkali protease from Bacillus used in washing powders
 XX
 PS Disclosure; Page 58-63; 71pp; Japanese.
 XX
 CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease.
 XX
 SO Sequence 640 AA;
 Query Match 97.5%; Score 3022; DB 20; Length 640;
 Best Local Similarity 93.3%; Pred. No. 1.9e-237;
 Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
 OY 2 RKKVFLSVLSAAALSTVVALXNPAGARXEDLDKGIQTDTDXGFSKXOTGAAPL 61
 DB 3 RKKVFLSVLSAAALSTVVALXNPAGARXEDLDKGIQTDTDXGFSKXOTGAAPL 62
 OY 62 VESENVKLKGLKKLETPANNKLIHQFNGPILEETKQXLEXTGAKLIDYIPYAYI 121
 DB 63 VESENVKLKGLKKLETPANNKLIHQFNGPILEETKQXLEXTGAKLIDYIPYAYI 122
 OY 122 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 181
 DB 123 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 182
 OY 182 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 241
 DB 183 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 242
 OY 242 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 301
 DB 243 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 302
 OY 302 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 361
 DB 303 VEREGVYKXXXXIEHVESVEPLPYXIDPOLFTGASXLYKAAALDTKXNKEVOLRG 362

Query Match 97.4%; Score 3021; DB 20; Length 640;

[illegible]

XX 07-OCT-1997: 97JP-0274570.
 PR (KAOS) KAO CORP.
 XX
 PA Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
 PI Saeki K, Shikata S, Takaiwa M;
 XX
 DR WPI: 1999-287736/24.
 N-PSDB: AAX37277.
 XX
 PT Alkali protease from Bacillus used in washing powders
 PS
 XX Claim 3; Page 47-50; 71pp; Japanese.
 CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidizing agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidizing agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease of the
 CC invention.
 XX
 XX Sequence 639 AA;
 S0
 Query Match 97.3%; Score 3017; DB 20; Length 639;
 Best Local Similarity 100.0%; Pred. No. 4, 8e-237;
 Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRKKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIQTITTDXXGFSKXOTGAAPF 60
 DB 1 MRKKVFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIQTITTDXXGFSKXOTGAAPF 60
 QY 61 LVESNNVLAAGLKKKLETPVANKKLIHXONGPILETKOXLEKTKAKIIDYIPDAYI 120
 DB 61 LVESNNVLAAGLKKKLETPVANKKLIHXONGPILETKOXLEKTKAKIIDYIPDAYI 120
 QY 121 VEYEGDVYXXXXXIEHVESVEPYLPYXIIDPOLFTKGASLVKAKALDTKXKNEVQLRG 180
 DB 121 VEYEGDVYXXXXXIEHVESVEPYLPYXIIDPOLFTKGASLVKAKALDTKXKNEVQLRG 180
 QY 181 IEXIAQXXXXNDVYXITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDT 240
 DB 181 IEXIAQXXXXNDVYXITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDT 240
 QY 241 GRNDSMHEARFKITAIYALGRNNANDTNGHGHVAGSVLGNKXTTKKGAPOANLVEQ 300
 DB 241 GRNDSMHEARFKITAIYALGRNNANDTNGHGHVAGSVLGNKXTTKKGAPOANLVEQ 300
 QY 301 SIMDSXGIGLGPSTLQTLFQOAXSAGARHTNSGAAVNGAYTTDSNNVDYVRKNDMT 360
 DB 301 SIMDSXGIGLGPSTLQTLFQOAXSAGARHTNSGAAVNGAYTTDSNNVDYVRKNDMT 360
 QY 361 ILPAAGNEXPMGDTISAFGTAKMAITVGATENLRPSFGSVADNINHVAQFSSRGFTKGR 420
 DB 361 ILPAAGNEXPMGDTISAFGTAKMAITVGATENLRPSFGSVADNINHVAQFSSRGFTKGR 420
 QY 421 IKPDVMAAGTYILSRSSLPADSSPMANHDSKYAVMGSTSWATPTVAANVQLRHFVK 480
 DB 421 IKPDVMAAGTYILSRSSLPADSSPMANHDSKYAVMGSTSWATPTVAANVQLRHFVK 480
 QY 481 RGIITPKPBLAKAALAGADGGLGYPNGOQWGRVTLDDKSLINVAIVNSSXLSSTOKATY 540
 DB 481 RGIITPKPBLAKAALAGADGGLGYPNGOQWGRVTLDDKSLINVAIVNSSXLSSTOKATY 540
 QY 541 XFTATAGKPLKISLVWSDAPASTASTVTLVNDLVLITAPNGTXYVGNDPXXPXXXXXNDWG 600

DB 541 XFTATAGKPLKISLVWSDAPASTASTVTLVNDLVLITAPNGTXYVGNDPXXPXXXXXNDWG 600
 QY 601 RNNVENWFINKPQSGTYTIEVOAYNVVPGPOXFEIAYN 639
 DB 601 RNNVENWFINKPQSGTYTIEVOAYNVVPGPOXFEIAYN 639
 RESULT 5
 AAY17088
 ID AAY17088 standard; protein; 640 AA.
 XX
 AC AAY17088;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE An alkaline protease sequence from Bacillus species.
 XX
 KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidizing agent.
 XX
 OS Bacillus sp.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..640
 FT N-PSDB: AAX37278.
 FT "all residues indicated as Xaa are arbitrary amino acids"
 XX
 PN W09918218-A1.
 XX
 PD 15-APR-1999.
 XX
 PE 07-OCT-1998; 98WO-JP04528.
 XX
 PR 07-OCT-1997; 97JP-0274570.
 XX
 XX (KAOS) KAO CORP.
 PA Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
 PI Saeki K, Shikata S, Takaiwa M;
 XX
 DR WPI: 1999-287736/24.
 N-PSDB: AAX37278.
 XX
 PT Alkali protease from Bacillus used in washing powders
 PS
 XX Claim 3; Page 50-53; 71pp; Japanese.
 CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidizing agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidizing agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease of the
 CC invention.
 XX
 XX Sequence 640 AA;
 S0
 Query Match 97.0%; Score 3007; DB 20; Length 640;
 Best Local Similarity 100.0%; Pred. No. 3, 1e-236;
 Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 KKRYFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIQTITTDXXGFSKXOTGAAPFV 62
 DB 4 KKRYFLSVLSAAALISTVALXNPSAGXARXFDLDFKGIQTITTDXXGFSKXOTGAAPFV 63

```

OY 63 ESENVKLAKGLKKLETPVANNKLIHXPNGPILEETKQLEXTGAKILDYIPDAVIVE 122
DB 64 ESENVKLAKGLKKLETPVANNKLIHXPNGPILEETKQLEXTGAKILDYIPDAVIVE 123
OY 123 YEGDVXKXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVAXALDTRKQNKKEVQKRGIE 182
DB 124 YEGDVXKXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVAXALDTRKQNKKEVQKRGIE 183
OY 183 XIAOXXXNDVXYITAKPEKVMNDVARGIVKADVAOSYGLYGOGIVAVAPGTGIDTGR 242
DB 184 XIAOXXXNDVXYITAKPEKVMNDVARGIVKADVAOSYGLYGOGIVAVAPGTGIDTGR 243
OY 243 NDSSMHEAFRGKITALYALGRTNNANDTNGHGTTHVAGSVLNGXTNKGAPQANLVFQSI 302
DB 244 NDSSMHEAFRGKITALYALGRTNNANDTNGHGTTHVAGSVLNGXTNKGAPQANLVFQSI 303
OY 303 MDSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKNDMTIL 362
DB 304 MDSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKNDMTIL 363
OY 363 FAAGNEXPNNGGTSAPGTAKNATVGTATENLRPSFGSYADNINHVAOFSSRGPTKDGRIK 422
DB 364 FAAGNEXPNNGGTSAPGTAKNATVGTATENLRPSFGSYADNINHVAOFSSRGPTKDGRIK 423
OY 423 PDVMAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQLREHFVKNG 482
DB 424 PDVMAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQLREHFVKNG 483
OY 483 ITPKPSLLKALINAGADGGLGYPNGNGMGRVTLDRKSLNAYVNESSXLSTQKATYTF 542
DB 484 ITPKPSLLKALINAGADGGLGYPNGNGMGRVTLDRKSLNAYVNESSXLSTQKATYTF 543
OY 543 TATGKPLKISLVWSDAPASTTASVTLVNDLDTVITANCGXYVGNPFXXXNMDGN 602
DB 544 TATGKPLKISLVWSDAPASTTASVTLVNDLDTVITANCGXYVGNPFXXXNMDGN 603
OY 603 NVENVFIXPOSGTYTTEVOAVNPVGPQESLAVN 639
DB 604 NVENVFIXPOSGTYTTEVOAVNPVGPQESLAVN 640

```

RESULT 6
AAW89547
ID AAW89547 standard; Protein; 641 AA.
AAW89547:
12-APR-1999 (first entry)
DE Bacillus jpl170 protease.
KW Protease; detergent; surfactant; leather processing; debittering;
flavour.
OS Bacillus sp.
FH Key Location/Qualifiers
FT Peptide 1..33
FT Region /note= "signal peptide"
FT /note= "34..208" prepro region"
FT Protein /note= "209..641"
FT /note= "mature protein"
PN W08856927-A2.
PD 17-DEC-1998.
PE 09-JUN-1998; 98MO-US12005.
PR 12-JUN-1997; 97US-0873479.
PA (NOVO) NOVO NORDISK BIOTECH INC.

```

XX Christianson L, Sloma A;
PI WPI: 1999-080908/07.
DR N-PSDB; AAW82382.
XX Novel protease from Bacillus subtilis 1620 - useful in laundry and
PT dishwashing detergents and for leather processing
XX Claim 7; Page 53-54; 77pp: English.
PS
CC This is the amino acid sequence of a novel protease of Bacillus sp.
CC jpl170 (NC18 12513), as deduced from the nucleotide sequence of an
CC isolated gene (see AAW82382). The entire protein, including of an
CC signal peptide and prepro region, has 77% identity to alkaline
CC protease Y (see AAW89548) from Bacillus.
CC The invention provides
CC vectors, recombinant host cells and methods for the recombinant
CC production of the protease. The protease is used in laundry and
CC dishwashing detergents, for institutional and industrial cleaning,
CC and for leather processing, as well as for debittering and
CC enhancing the degree of hydrolysis of protein hydrolysates, for
CC flavour development through hydrolysis of proteins, degradation of
CC undesired peptides and in enzymatic synthesis of peptides. It has
CC enhanced stability towards oxidation under alkaline conditions,
CC e.g. towards bleaching agents of the peroxy type. The invention
CC also provides mutant cells in which the protease activity is
CC diminished. Such cells can be used for the production of
CC heterologous recombinant proteins.
SO Sequence 641 AA:

```

Query Match 87.8%; Score 2723; DB 20; Length 641;
Best Local Similarity 82.9%; Pred. No. 4,1e-213;
Matches 532; Conservative 40; Mismatches 66; Indels 4; Gaps 2;

```

OY 1 MKK---KVFSLVLSAAAILSTVALNPSAGAXKRPDIDFKGIQT7TDXXGFSKQQTGA 57
DB 1 MKKSKKRVFLSVLSVALLSVALSPSTGANNFELDFGIELTLEKAAKRGCKTG 60
OY 58 AAFVSENVKLKXGLKKLETPVANNKLIHXPNGPILEETKQLEXTGAKILDYIPY 117
DB 61 ASFLVSENVKIPKSIQKLEVPADNKLIVQFDGPILIEFTQLQLEGTAKILDYIPY 120
OY 118 AYIVEEGDVXKXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVAXALDTRKQNKKEVO 177
DB 121 AYIVEEGDVXKXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVAXALDTRKQNKKEVO 180
OY 178 LRGIXIXIAOXXXNDVXYITAKPEKVMNDVARGIVKADVAOSYGLYGOGIVAVAPGTG 237
DB 181 LRGLEIOIAQVATFNNDDVLYTPKPEYELVDVARGIVKADVAONNPFGLGCGQIVAAVDG 240
OY 238 LDTGRNDSMHEAFRGKITALYALGRTNNANDTNGHGTTHVAGSVLNGXTNKGAPQANLV 297
DB 241 LDTGRNDSMHEAFRGKITALYALGRTNNANDTNGHGTTHVAGSVLNGXTNKGAPQANLV 299
OY 298 VFQSTIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKN 357
DB 300 VFQSTIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAANVNGAYTTDSRVNDYVRKN 359
OY 358 DMTILFAAGNEXPNNGGTSAPGTAKNATVGTATENLRPSFGSYADNINHVAOFSSRGPTK 417
DB 360 DMTILFAAGNEXPNNGGTSAPGTAKNATVGTATENLRPSFGSYADNINHVAOFSSRGPTK 419
OY 418 DGRIRPDVMAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQLREHF 477
DB 420 DGRIRPDVMAFGTXILSARSLAPDSFMANHDSKYAYMGTSMATPIYAGVNAQLREHF 479
OY 478 VKNRGITPKPSLLKALINAGADGGLGYPNGNGMGRVTLDRKSLNAYVNESSXLSTQK 537
DB 480 VKNRGVTPKPSLLKALINAGADGGLGYPNGNGMGRVTLDRKSLNAYVNESSXLSTQK 539
OY 538 ATYXFTATAGKPLKISLVWSDAPASTTASVTLVNDLDTVITANCGXYVGNPFXXXNMDGN 597

```

Db 540 ATYSEFAQAGKPLKISLWSDAPGSTTASTLTVDLVTAPNGTRYVGNDFAPYDNN 599
 QY 598 WGRNNVENVFTNPGSGTYTEQAYNVPQXSLAIYN 639
 Db 600 WGRNNVENVFTNPGSGTYTEQAYNVPSPQTFSLAIHV 641

RESULT 7

AAM89548
 ID AAM89548 standard: Protein: 636 AA.

AC AAM89548;

DT 12-APR-1999 (first entry)

DE Bacillus sp. alkaline protease Y.

KW Alkaline protease Y; detergent; surfactant; leather processing;

KM debittering; flavour.

OS Bacillus sp.

PM WO9856927-A2.

PD 17-DEC-1998.

PF 09-JUN-1998; 98MO-US12005.

PR 12-JUN-1997; 97US-0873479.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Christianson L, Sloma A;

XX WPI, 1999-080908/07.

PT Novel protease from Bacillus subtilis LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing

PS Claim 3; Page 55-56; 77pp; English.

CC This is the amino acid sequence of a Bacillus sp. alkaline protease
 CC Y that is said to have good alkali and surfactant resistance and
 CC improved detergency. It shows 77% identity to a newly isolated
 CC protease (see AAM89547) of Bacillus sp. jpl10 (NCIF 12513). The
 CC invention provides vectors, recombinant host cells and methods for
 CC the recombinant production of such proteases. The protease are
 CC used in laundry and dishwashing detergents, for institutional and
 CC industrial cleaning, and for leather processing, as well as for
 CC debittering and enhancing the degree of hydrolysis of protein
 CC hydrolysates, for flavour development through hydrolysis of
 CC proteins, degradation of undesired peptides and in enzymatic
 CC synthesis of peptides. They have enhanced stability towards
 CC oxidation under alkaline conditions, e.g. towards bleaching agents
 CC of the peroxy type. The invention also provides mutant cells in
 CC which the protease activity is diminished. Such cells can be used
 CC for the production of heterologous recombinant proteins.

CC Sequence 636 AA;

Query Match 77.7%; Score 2408; DB 20; Length 636;

Best Local Similarity 73.2%; Pred. No. 1,8e-187;

Matches 467; Conservative 67; Mismatches 98; Indels 6; Gaps 5;

QY 3 KKKVFLSVLSAAAILSTVALNKPASGAXKXFDLDFKGIQTTDXGFSKQXGTGAAPLV 62
 Db 4 KRRVLSVYASAAIILASVSVSPTSGA--DFQVFNQGVK-SLEMSLVKPISSGASFLV 60
 QY 63 ESENKIKKGLKKLETPYANKKLHIQFN-GPILEETQXILEXGAKLIDYIPYATYV 121
 Db 61 DTEININIPGIDQKLEAVQKDNELIYQFTPGPISEEEKRSGSLGVSTLDYVPPDAFV 120
 QY 122 EYEGDVXSKXXXIHEHVESVEPYLPXYXIDPQLFTGASGLVAXXALDTPQXKKEVQLRGI 181

Db 121 QYSG-ATKNISLHSEVENYQPFLLPKYIDPELLTGASQLVQAVILNTKHEKNKKFTGL 179
 QY 182 EXIAQXXXNDVXYITTAKEEYKVMNDVANGYKADVAOSSYGLGCGQIVAAVDGLDTG 241
 Db 180 DEIYQAAANNNDLYISPKPEYEELMNDVANGYKADVAQNNYGLGCGQLVAAVDGLDTG 239
 QY 242 RNDSSWHEAFRGKITALYALVGRTNNAANDFNHGHGTHVAGSVLGNKGMQAQANLVPOS 301
 Db 240 RNDSSWHEAFRGKITALYALVGRTNNAANDFNHGHGTHVAGSVLGNKGMQAQANLVPOS 298
 QY 302 IMDSXGGLGGLPSNLQTLFSAQXASAGARLHTNSWGAAYVAYTTDSRNVDDYVRKNDMTI 361
 Db 299 IMDSXGGLGGLPSNLQTLFSAQXAMNAGARLHTNSWGAAYVAYTTDSRNVDDYVRKNDMTI 358
 QY 362 LFAAGNEXNGGTISAPGAKNAIYVATENLRPSGSIADININVAQSSSGPTKDGRI 421
 Db 359 LFAAGNEXNGGTISAPGAKNAIYVATENLRPSGSIADININVAQSSSGPTKDGRI 418
 QY 422 KPDVMAFGTXXLSARSLAPDSSFFMANHDSKYAYMGCTSMATPIYAGNVAQLREHFVKR 481
 Db 419 KPDVMAFGTXXLSARSLAPDSSFFMANHDSKYAYMGCTSMATPIYAGNVAQLREHFVKR 478
 QY 482 GITPKRSLKALALAGADXLGYPNGMGWGRVTLDKSLNAYVNESSXLSTSQATYX 541
 Db 479 GITPKRSLKALALAGADVLGYPSDGGWGRVTLDKSLNAYVNEATALATGOKATYX 538
 QY 542 FTATAGKPLKISLWSDAPGSTTASTLTVDLVTAPNGTRYVGNDFAPYDNN 601
 Db 539 FOAAGKPLKISLWSDAPGSTTASTLTVDLVTAPNGTRYVGNDFAPYDNN 598
 QY 602 NNVENVFINKPGSGTYTEQAYNVPQXSLAIYN 639
 Db 599 NNVENVFINKPGSGTYTEQAYNVPQXSLAIYN 636

RESULT 8
 AAM50080
 ID AAM50080 standard: Protein: 434 AA.

AC AAM50080;

DT 12-AUG-2002 (first entry)

DE Bacillus sp KSM-KP43 alkaline protease protein fragment.

KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

OS Bacillus sp.

PM EP1209233-A2.

PD 29-MAY-2002.

PF 22-NOV-2001; 2001EP-0127851.

PR 22-NOV-2000; 2000JP-0355166.

PR 12-APR-2001; 2001JP-0114048.

XX (KAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

XX Okuda M, Saeki K;

XX WPI, 2002-437518/47.

PS Claim 1; Page 10-11; 25pp; English.

CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &

CC (34 - 38*) compared to prior art alkaline proteases (31 and 23*). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC Bacillus sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50090.
 XX
 SQ Sequence 434 AA:
 Query Match 69.5%; Score 2155; DB 23; Length 434;
 Best Local Similarity 96.3%; Pred. No. 4.3e-167;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 OY 206 NDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLYALGRTN 265
 DB 1 NDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLYALGRTN 60
 OY 266 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 325
 DB 61 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 120
 OY 326 AGARHTNSMGAALVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
 DB 121 AGARHTNSMGAALVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
 OY 386 TVGATENLRPFSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSF 445
 DB 181 TVGATENLRPFSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSF 240
 OY 446 WANHDSKYAAMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSILKAALLAGAADGIGY 505
 DB 241 WANHDSKYAAMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSILKAALLAGAADGIGY 300
 OY 506 PNGNQGGRVTLDKSLNVAIVYNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 565
 DB 301 PNGNQGGRVTLDKSLNVAIVYNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 360
 OY 566 SVTLVNDLDLVITAPNGTYVGNDFXEXXKXNMDCRRNNEVFINKPQSGTYTIEVOAYN 625
 DB 361 SVTLVNDLDLVITAPNGTYVGNDFXEXXKXNMDCRRNNEVFINKPQSGTYTIEVOAYN 420
 OY 626 VPVGPOQXSLAIYN 639
 DB 421 VPVGPOQXSLAIYN 434
 RESULT 9
 AAM50081
 ID AAM50081 standard; protein: 434 AA.
 XX
 AC AAM50081;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Bacillus sp KSM-KP9860 alkaline protease protein fragment.
 XX
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS Bacillus sp.
 XX
 PN EPI209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-0127851.
 XX
 PR 22-NOV-2000; 2000JP-0355166.
 PR 12-APR-2001; 2001JP-0114048.
 XX
 PA (KAO) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Arai H, Sumitomo N;
 PI Okuda M, Saeki K;
 DR WPI: 2002-437518/47.

XX
 PT New modified alkaline proteases useful in detergent compositions -
 XX
 PS Claim 5; Page 12-13; 25pp; English.
 XX
 CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38*) compared to prior art alkaline proteases (31 and 23*). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC Bacillus sp strain KSM-KP9860 described in the method of the invention.
 XX
 SQ Sequence 434 AA:
 Query Match 69.5%; Score 2155; DB 23; Length 434;
 Best Local Similarity 96.3%; Pred. No. 4.3e-167;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 OY 206 NDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLYALGRTN 265
 DB 1 NDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITLYALGRTN 60
 OY 266 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 325
 DB 61 NANTNGHGHVAGSVLGNKGNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAXS 120
 OY 326 AGARHTNSMGAALVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
 DB 121 AGARHTNSMGAALVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
 OY 386 TVGATENLRPFSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSF 445
 DB 181 TVGATENLRPFSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSF 240
 OY 446 WANHDSKYAAMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSILKAALLAGAADGIGY 505
 DB 241 WANHDSKYAAMGCTSMATPIVAGNVAOLREHFVKNRGITPKPSILKAALLAGAADGIGY 300
 OY 506 PNGNQGGRVTLDKSLNVAIVYNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 565
 DB 301 PNGNQGGRVTLDKSLNVAIVYNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 360
 OY 566 SVTLVNDLDLVITAPNGTYVGNDFXEXXKXNMDCRRNNEVFINKPQSGTYTIEVOAYN 625
 DB 361 SVTLVNDLDLVITAPNGTYVGNDFXEXXKXNMDCRRNNEVFINKPQSGTYTIEVOAYN 420
 OY 626 VPVGPOQXSLAIYN 639
 DB 421 VPVGPOQXSLAIYN 434
 RESULT 10
 AAM50085
 ID AAM50085 standard; protein: 434 AA.
 XX
 AC AAM50085;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Bacillus sp alkaline protease protein A-1 fragment.
 XX
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS Bacillus sp.
 XX
 PN EPI209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-0127851.
 XX
 PR 22-NOV-2000; 2000JP-0355166.
 PR 12-APR-2001; 2001JP-0114048.

XX (KAOS) KAO CORP.
 XX PA Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX PI Okuda M, Saeki K;
 XX DR WPI: 2002-437518/47.
 XX
 XX New modified alkaline proteases useful in detergent compositions -
 XX
 XX PS Claim 5; Page 18-19; 25pp; English.
 XX
 XX This invention describes novel *Bacillus* sp. alkaline proteases useful in
 XX detergent compositions, especially in laundry, bleaching or automatic
 XX dishwasher detergents. The novel proteases have an increased detergency %
 XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 XX sequence represents a fragment of the alkaline protease A-1 from
 XX *Bacillus* sp NCIB12289 described in the method of the invention.
 XX
 XX SQ Sequence 434 AA;

Query Match 67.1%; Score 2082; DB 23; Length 434;
 Best Local Similarity 91.5%; Pred. No. 3,8e-161;
 Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 206 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 265
 DB 1 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 60
 QY 266 NANDTNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAYS 325
 DB 61 NANDPNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAYS 120
 QY 326 AGARITHTNSMGAAYNCATYTDSSRNDDYVRKNDMTLFAAGNEXNGRTISAPGAKNAI 385
 DB 121 AGARITHTNSMGAAYNCATYTDSSRNDDYVRKNDMTLFAAGNEXNGRTISAPGAKNAI 180
 QY 386 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 445
 DB 181 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 240
 QY 446 WANHDSKYAYMGTSMAPTIVAGNVAQLREHFVKNRGITPKPSLLKAAALAGADVGLGY 505
 DB 241 WANHDSKYAYMGTSMAPTIVAGNVAQLREHFVKNRGITPKPSLLKAAALAGADVGLGY 300
 QY 506 PNGNOGMRVTLDKSLNVAIVYNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 565
 DB 301 PNGNOGMRVTLDKSLNVAIVYNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 360
 QY 566 SVTLVNDLVLITAPNGTYYVGNDEXXPYXXNMWDRNNVENVEFINXPOSGTYTIEVOAYN 625
 DB 361 SVTLVNDLVLITAPNGTYYVGNDEXXPYXXNMWDRNNVENVEFINXPOSGTYTIEVOAYN 420
 QY 626 VEVGPQXFSALIVN 639
 DB 421 VEVGPQXFSALIVN 434

RESULT 11
 AAM50086
 ID AAM50086 standard; protein; 433 AA.
 XX
 XX AAM50086;
 XX
 XX DT 12-AUG-2002 (first entry)
 XX
 XX DE *Bacillus* sp alkaline protease protein A-2 fragment.
 XX
 XX KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 XX OS *Bacillus* sp.
 XX
 XX PN EP1209233-A2.

XX PD 29-MAY-2002.
 XX PF 22-NOV-2001; 2001EP-0127851.
 XX PR 22-NOV-2000; 2000JP-0355166.
 XX PR 12-APR-2001; 2001JP-0114048.
 XX
 XX PA (KAOS) KAO CORP.
 XX
 XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX PI Okuda M, Saeki K;
 XX DR WPI: 2002-437518/47.
 XX
 XX New modified alkaline proteases useful in detergent compositions -
 XX
 XX PS Claim 5; Page 20-21; 25pp; English.
 XX
 XX This invention describes novel *Bacillus* sp. alkaline proteases useful in
 XX detergent compositions, especially in laundry, bleaching or automatic
 XX dishwasher detergents. The novel proteases have an increased detergency %
 XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 XX sequence represents a fragment of the alkaline protease A-2 from
 XX *Bacillus* sp NCIB12513 described in the method of the invention.
 XX
 XX SQ Sequence 433 AA;

Query Match 66.4%; Score 2060.5; DB 23; Length 433;
 Best Local Similarity 91.7%; Pred. No. 2.1e-159;
 Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;

QY 206 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 265
 DB 1 NDVARGIVADVAOSSYGLYGOGIYAVADTGLDGRNDSMHEAFRGKITALYALGRIN 60
 QY 266 NANDTNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAYS 325
 DB 61 NANDPNGHGTTHVAGSVLNGXTNKGAPQANLVFQSIMDSXGGLGGLPSNLQTLFSQAYS 119
 QY 326 AGARITHTNSMGAAYNCATYTDSSRNDDYVRKNDMTLFAAGNEXNGRTISAPGAKNAI 385
 DB 120 AGARITHTNSMGAAYNCATYTDSSRNDDYVRKNDMTLFAAGNEXNGRTISAPGAKNAI 179
 QY 386 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 445
 DB 180 TVGATENTLRPSFGSYADNINHAOFSSRGPTKDGRIKPDVMAPGTYILSARSILAPDSSF 239
 QY 446 WANHDSKYAYMGTSMAPTIVAGNVAQLREHFVKNRGITPKPSLLKAAALAGADVGLGY 505
 DB 240 WANHDSKYAYMGTSMAPTIVAGNVAQLREHFVKNRGITPKPSLLKAAALAGADVGLGY 299
 QY 506 PNGNOGMRVTLDKSLNVAIVYNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 565
 DB 300 PNGNOGMRVTLDKSLNVAIVYNESSXLSOKATYXFTTACKPLKISLWSDAPASTA 359
 QY 566 SVTLVNDLVLITAPNGTYYVGNDEXXPYXXNMWDRNNVENVEFINXPOSGTYTIEVOAYN 625
 DB 360 SVTLVNDLVLITAPNGTYYVGNDEXXPYXXNMWDRNNVENVEFINXPOSGTYTIEVOAYN 419
 QY 626 VEVGPQXFSALIVN 639
 DB 420 VEVGPQXFSALIVN 433

RESULT 12
 AAM50090
 ID AAM50090 standard; protein; 434 AA.
 XX
 XX AAM50090;
 XX
 XX DT 12-AUG-2002 (first entry)
 XX

XX	DE	Bacillus sp KSM-KP43 alkaline protease protein variant.
KM	XX	alkaline protease; detergent; laundry; bleaching; dishwasher; mutant;
KW	XX	metin.
OS	OS	Bacillus sp.
XX	XX	synthetic.
PH	Key	Location/Qualifiers
FT	FT	Misc-difference 46
FT	FT	/label= 'y,w,a,d,e,t,v,l,i,h,s,k,g,m,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 54
FT	FT	/label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 57
FT	FT	/label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 66
FT	FT	/label= 'e,d,s,g,a,t,l,m,c,v,g,i
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 84
FT	FT	/label= OTHER, R
FT	FT	/note= "OTHER- deleted residue. Specifically described in Claim 1"
FT	FT	Misc-difference 101..106
FT	FT	/label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 104
FT	FT	/label= OTHER, P
FT	FT	/note= "OTHER- deleted residue. Specifically described in Claim 1"
FT	FT	Misc-difference 107
FT	FT	/label= 'k,r,a,s
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 119
FT	FT	/label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 124
FT	FT	/label= 'a,k
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 136
FT	FT	/label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 138
FT	FT	/label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 148
FT	FT	/label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 193
FT	FT	/label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 195
FT	FT	/label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 205
FT	FT	/label= 'y,w,a,d,e,t,v,l,i,h,s,k,q,m,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 247
FT	FT	/label= 'w,f,a,r,e,t,v,l,i,h,s,e,m,c
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 256
FT	FT	/label= OTHER, a,s,e,v,l,r,e,d
FT	FT	/note= "OTHER- deleted residue. Specifically described in Claim 1"
FT	FT	Misc-difference 264
FT	FT	/label= 'e,d,s,g,a,t,l,m,c,v,g,i
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 257
FT	FT	/label= 'v,i
FT	FT	/note= 'as claimed in Claim 3'
FT	FT	Misc-difference 342
FT	FT	/note= 'as claimed in Claim 3'

FT	/label= k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
FT	/note= "as claimed in Claim 3"
FT	Misc-difference 369
FT	/label= OTHER, d
FT	/note= "OTHER- deleted residue. Specifically described
FT	in Claim 1"
EN	
PN	EP1209233-A2.
XX	
ED	29-MAY-2002.
XX	
KX	22-NOV-2001; 2001EP-0127851.
PF	
PR	22-NOV-2000; 2000JP-0355166.
PR	12-APR-2001; 2001JP-0114048.
PA	(KAOS) KAO CORP.
XX	
P1	Hatada Y., Ogawa A., Kageyama Y., Sato T., Ataki H., Sumitomo N;
P1	Okuda M., Saeki K;
XX	
XX	WPI; 2002-437518/47.
PT	
PT	New modified alkaline proteases useful in detergent compositions -
XX	
XX	Claim 1; Page -: 25pp; English.
CC	This invention describes novel Bacillus sp. alkaline proteases useful in
CC	detergent compositions, especially in laundry, bleaching or automatic
CC	dishwasher detergents. The novel proteases have an increased detergency &
CC	(34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
CC	sequence represents a variant of the alkaline protease KP43 from
CC	Bacillus sp strain KSM-KP43 created from the wild-type protease
CC	represented in AA050080
CC	NOTE: This sequence is not represented in the specification but has
CC	been constructed from the sequence represented in record AA050080.
XX	
XX	Sequence 434 AA:
QY	Query Match 64.3%; Score 1994; DB 23; Length 434:
	Best Local Similarity 91.2%; Pred. No. 5.5e-154;
	Matches 336; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY	206 NDVARGIVKADVAOSSYGLYGOGQIYAADVDTGDLTGRNDSMHEAFRCGTALYALGRTN 265
Db	1 NDVARGIVKADVAOSSYGLYGOGQIYAADVDTGDLTGRNDSMHEAFRCGTALYALKRTN 60
QY	266 NANDTIGHGHTHAGSVLGNCGTXRKKGAPQANLVFPOSIMDXSGCLGSLPEINLOTQTFSOAXS 325
Db	61 NANDTIGHGHTHAGSVLGNGSTNXGMAPQANLVFPOSIMDXSGXXXXXPENLOTQTFSOAXS 120
QY	326 AGARITNTSMGAANGAYTTDSRNVDVYRKNDMTILFAAGNXPXPGCTISAPCAKNAI 385
Db	121 AGAXIHITNSMGAAVNAXXTTDSRNVDVYRKNDMTILFAAGNCPGCGTISAPGAKNAI 180
QY	386 TYVATFNTRPSFGSYADININHVAQFSRRPKDGRIRPDVMAPGCTIISASSSLAPDSSF 445
Db	181 TYVATEFNLRPSFXSXADNIINHVAQXSRRPKDGRIRPDVMACGCTIISARSLSLAPDSSF 240
QY	446 WANHSKXAAYKGGCSMATPIVACNVNOLREHYVNRGRTETKPSLKAAALAGADXGXY 505
Db	241 WANHSXAIAKMGSSXXTPRIYACXVAQLREHYVNRKRGTIKPSLLAALLAGADIGLGY 300
QY	506 PNCNGCMGRVTLDKSLNAVAVNESSXLSTSOKATYXTTATAGAPFLKISLWSDPASTTA 565
Db	301 PNCNGCMGVTLTDKSLNAVAVNESSLSLTSOKATYSFTATAKPKLTISLWSMDPASTTA 360
QY	566 SVTVLVNDLDELVTAPNGTXYYVGNDEFPXPKXNMDCRRNNVENVEFINNPQSQTYYIEQAVN 625
Db	361 SVTVLVNDLKLVTAPNGTOYVGNDFTSPYNNDMMGCRNNVENVEFINNPQSQTYYIEQAVN 420
QY	626 VVPGPQXFESLAIVN 639

Db 421 VPGPQFSLAIVN 434

RESULT 13
AAM50084
ID AAM50084 standard; protein: 433 AA.
XX
XX AAM50084;
AC
XX 12-AUG-2002 (first entry)
XX
DE Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.
XX
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
XX Bacillus sp.
XX OS EPI209233-A2.
XX PN 29-MAY-2002.
XX PD 22-NOV-2001; 2001EP-0127851.
XX PF 22-NOV-2001; 2000JP-0355166.
XX PR 12-APR-2001; 2001JP-0114048.
XX
XX (KAOS) KAO CORP.
XX
XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeki K;
XX WPI: 2002-437518/47.
XX
PT New modified alkaline proteases useful in detergent compositions -
XX
XX Claim 5; Page 16-18; 25pp; English.
XX
XX This invention describes novel Bacillus sp. alkaline proteases useful in
CC detergent compositions, especially in laundry, bleaching or automatic
CC dishwasher detergents. The novel proteases have an increased detergency &
CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
CC sequence represents a fragment of the alkaline protease SD-521 from
CC Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the
CC invention.
XX
XX Sequence 433 AA;
SQ

Query Match 63.0%; Score 1952.5; DB 23; Length 433;
Best Local Similarity 86.6%; Pred. No. 1.3e-150;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 206 NDVARGIVKADYAOSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITALTALVAGRTN 265
DB 1 NDVARGIVKADYAOSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITALTALVAGRTN 60

QY 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSIMDSXGGLGSPNLQTLFSSQAXS 325
DB 61 NANDNGHGTTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSXGGLGSPNLQTLFSSQAMN 119

QY 326 AGARHTHTSMGCAAVNGATTTDSRNDDVVRKNDMTILFAAGNEXNGGTISAPGTAKNAI 385
DB 120 AGARHTHTSMGCAVPVNGATYANSROYDEVYRNDMTVLFAAGNEGPNSGTISAPGTAKNAI 179

QY 386 TVGATEENLRPSFGSYADININHYAOFSSRGPTKGRIPDVMAPTXIIISARSSLAPDSSF 445
DB 180 TVGATEENLRPSFGSLADPNPHIAOFSRGATRDGRIKPDYVATPGFTIISARSSLAPDSSF 239

QY 446 WANHDSKYAYMGTSMAPIVAGNAVOLREHEVKNRGITPKPSLKAALIAAGADGXIGY 505
DB 240 WANYSKYAYMGTSMAPIVAGNAVOLREHEIKNRGITPKPSLKAALIAAGATVGIGY 299

QY 506 PNGNGMGRTVILDKSLNAYVNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTTA 565
I:|||||

Db 300 PSGDQMGCRVTLDKSLNAYVNEATALATGQATYSFOAQAGKPLKISLWTDAGSTTA 359

QY 566 SVTLVNDLDTVTAPNGTYVGNDFXXPXXMMDSGRNNVENFINXPOSGTYTIEVOAYN 625
I:|||||

Db 360 SYTLVNDLDTVTAPNGQYVGNDFSYPDNMDGRNNVENFINAPSGTYTIEVOAYN 419

QY 626 VPGPQFSLAIVN 639
DB 420 VPGPQFSLAIVH 433

RESULT 14
AAM50082
ID AAM50082 standard; protein: 433 AA.
XX
XX AAM50082;
AC
XX 12-AUG-2002 (first entry)
XX
DE Bacillus sp D6 (FERM PI592) alkaline protease protein fragment.
XX
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
XX Bacillus sp.
XX OS EPI209233-A2.
XX PN 29-MAY-2002.
XX PD 22-NOV-2001; 2001EP-0127851.
XX PF 22-NOV-2001; 2000JP-0355166.
XX PR 12-APR-2001; 2001JP-0114048.
XX
XX (KAOS) KAO CORP.
XX
XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeki K;
XX WPI: 2002-437518/47.
XX
PT New modified alkaline proteases useful in detergent compositions -
XX
XX Claim 5; Page 13-15; 25pp; English.
XX
XX This invention describes novel Bacillus sp. alkaline proteases useful in
CC detergent compositions, especially in laundry, bleaching or automatic
CC dishwasher detergents. The novel proteases have an increased detergency &
CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
CC sequence represents a fragment of the alkaline protease E-1 from
CC Bacillus sp strain D6 (FERM-PI592) described in the method of the
CC invention.
XX
XX Sequence 433 AA;
SQ

Query Match 62.8%; Score 1948.5; DB 23; Length 433;
Best Local Similarity 86.4%; Pred. No. 2.8e-150;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 206 NDVARGIVKADYAOSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITALTALVAGRTN 265
DB 1 NDVARGIVKADYAOSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITALTALVAGRTN 60

QY 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFQSIMDSXGGLGSPNLQTLFSSQAXS 325
DB 61 NANDNGHGTTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSXGGLGSPNLQTLFSSQAMN 119

QY 326 AGARHTHTSMGCAAVNGATTTDSRNDDVVRKNDMTILFAAGNEXNGGTISAPGTAKNAI 385
DB 120 AGARHTHTSMGCAVPVNGATYANSROYDEVYRNDMTVLFAAGNEGPNSGTISAPGTAKNAI 179

QY 386 TVGATEENLRPSFGSYADININHYAOFSSRGPTKGRIPDVMAPTXIIISARSSLAPDSSF 445
I:|||||

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:37:09 : Search time 13.9891 Seconds
(without alignments)
1343.997 Million cell updates/sec

Title: US-09-920-954-1

Sequence: 1 MRKKVFLSVLSAAALSTV.....EVOATNVPGVQXPSIAIVN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/aa/PT03.COMB.pep:*
6: /cgn2_6/ptodata/1/aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3030	97.7	639	US-09-509-814A-4	Sequence 4, App1
2	3022	97.5	640	US-09-509-814A-6	Sequence 6, App1
3	3021	97.4	640	US-09-509-814A-8	Sequence 8, App1
4	3017	97.3	639	US-09-509-814A-1	Sequence 1, App1
5	3007	97.0	640	US-09-509-814A-2	Sequence 2, App1
6	2723	87.8	641	US-08-873-479-42	Sequence 42, App1
7	2417.5	78.0	635	US-08-873-479-43	Sequence 43, App1
8	1940.5	62.6	433	US-09-104-623A-4	Sequence 4, App1
9	1940.5	62.6	433	US-09-019-532-4	Sequence 1, App1
10	451.5	14.6	659	US-08-894-818B-1	Sequence 1, App1
11	451.5	14.6	659	US-09-445-472-12	Sequence 12, App1
12	408	13.2	654	US-08-894-818B-35	Sequence 15, App1
13	408	13.2	654	US-09-445-472-16	Sequence 16, App1
14	398	12.8	659	US-08-894-818B-5	Sequence 5, App1
15	391	12.6	412	US-09-445-472-1	Sequence 1, App1
16	391	12.6	522	US-08-894-818B-3	Sequence 3, App1
17	391	12.6	522	US-09-445-472-2	Sequence 4, App1
18	340	11.0	520	US-09-000-016-7	Sequence 7, App1
19	340	11.0	520	US-09-514-340-7	Sequence 4, App1
20	340	11.0	734	US-09-000-016-4	Sequence 4, App1
21	340	11.0	734	US-09-514-340-4	Sequence 2, App1
22	340	11.0	823	US-09-000-016-2	Sequence 2, App1
23	340	11.0	823	US-09-514-340-2	Sequence 1, App1
24	299.5	9.7	903	US-08-750-532-1	Sequence 1, App1
25	299.5	9.7	1398	US-08-750-532-9	Sequence 9, App1
26	299.5	9.7	1398	US-08-894-818B-8	Sequence 8, App1
27	299.5	9.7	1398	US-09-445-472-6	Sequence 6, App1

28	288.5	9.3	237	1	US-08-750-532-18	Sequence 18, App1
29	282	9.1	418	2	US-08-873-479-44	Sequence 44, App1
30	270.5	8.7	397	1	US-08-434-255-2	Sequence 2, App1
31	270.5	8.7	397	1	US-08-434-255-4	Sequence 4, App1
32	270.5	8.7	397	1	US-08-459-967-2	Sequence 2, App1
33	270.5	8.7	397	1	US-08-459-967-4	Sequence 4, App1
34	270.5	8.7	397	1	US-08-460-327-2	Sequence 2, App1
35	270.5	8.7	397	1	US-08-460-327-4	Sequence 4, App1
36	270.5	8.7	397	1	US-08-459-871-2	Sequence 2, App1
37	270.5	8.7	397	1	US-08-459-871-4	Sequence 4, App1
38	264	8.5	370	1	US-08-434-255-6	Sequence 6, App1
39	264	8.5	370	1	US-08-459-967-6	Sequence 6, App1
40	264	8.5	370	1	US-08-460-327-6	Sequence 6, App1
41	264	8.5	370	1	US-08-459-871-6	Sequence 6, App1
42	255	8.2	280	1	US-08-434-255-8	Sequence 8, App1
43	255	8.2	280	1	US-08-459-967-8	Sequence 8, App1
44	255	8.2	280	1	US-08-460-327-8	Sequence 8, App1
45	255	8.2	280	1	US-08-459-871-8	Sequence 8, App1

ALIGNMENTS

```
RESULT 1
US-09-509-814A-4
Sequence 4, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKATA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKET, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509, 814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 639
TYPE: PRP
ORGANISM: Bacillus sp.
US-09-509-814A-4
Query Match 97.7% Score 3030; DB 4; Length 639;
Best Local Similarity 93.4% Pred. No. 1.9e-254;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
OY 1 MRKKVFLSVLSAAALSTVALXNPSAGXARXFDLDFKGIOTTTDXGFSKXOXOTGAANF 60
DB 1 MRKKVFLSVLSAAALSTVALXNPSAGDARTFDLDFKGIOTTTDVSGFSKOROTGAANF 60
OY 61 LVESENKTLKGLKLETPVANNKLIHXFNGPILBETKXOLEXTGAILDIYIDYAYI 120
DB 61 LVESENKTLKGLKLETPVANNKLIHXFNGPILBETKXOLEXTGAILDIYIDYAYI 120
OY 121 VEYEGDVXSKXXIIEVESVEPYLPXYIIDPOLFTKGSXLYVAXMLDTKXKNKRVOLRG 180
DB 121 VEYEGDVXSKXXIIEVESVEPYLPXYIIDPOLFTKGSXLYVAXMLDTKXKNKRVOLRG 180
OY 181 IETXAOXXSXNDVXYTARPEYKVMNDVARGIYADVAOSSYGLYGGOGIYAVVDTGLDT 240
DB 181 IETXAOXXSXNDVXYTARPEYKVMNDVARGIYADVAOSSYGLYGGOGIYAVVDTGLDT 240
OY 241 GRNDSMHEAFROKITALYALGRTNNANDTNGHGTIVAGSVLGNCTXNKGAAPQANLVFO 300
```

```
|||||
Db 241 GNDSSMHEAFRGKITALTALGRTNNANDTNGHGVAGSVLGNGATNGMAPOAALVFO 300
QY 301 STMDSSGGLGGLPSNLQTLFSAQXSAGARIHTNSMGAAVNGATYTDSSRVNDVYRRNDMT 360
Db 301 STMDSSGGLGGLPSNLQTLFSAQXSAGARIHTNSMGAAVNGATYTDSSRVNDVYRRNDMT 360
QY 361 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 420
Db 361 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 420
QY 421 IMPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVVAQLREHFVNR 480
Db 421 IMPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVVAQLREHFVNR 480
QY 481 RGTTPPSLLKKAALJAGADGXGLGYPNGNGMGRTVLDKSLNVAAYNESSXLSSTQKATY 540
Db 481 RGTTPPSLLKKAALJAGADGXGLGYPNGNGMGRTVLDKSLNVAAYNESSXLSSTQKATY 540
QY 541 XFTATGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXPXKXNMWDG 600
Db 541 XFTATGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXPXKXNMWDG 600
QY 601 RNNVENVFINXPOSGTYTTEVOAYNVPGPQFSLAIYN 639
Db 601 RNNVENVFINXPOSGTYTTEVOAYNVPGPQFSLAIYN 639
```

RESULT 2

```
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
```

```
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6
```

```
Query Match          97.5%; Score 3022; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 9.3e-254;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 2 RKKKVFSLVSAALSTVALANPSAGXARXFDLDFKGIQTITDXXGFSKQXOTGAALFL 61
Db 3 KKKKVFSLVSAALSTVALANPSAGXARXFDLDFKGIQTITDXXGFSKQXOTGAALFL 62
QY 62 VESENVKLLKGLKKLETVPANNNKLHIYQNGPILLETQKXILEXTGAKILDIYIPYAYIV 121
Db 63 VESENVKLLKGLKKLETVPANNNKLHIYQNGPILLETQKXILEXTGAKILDIYIPYAYIV 122
QY 122 ESEGVVXSXXXIIEHVESVEPLPYXIIDPOLFTKGASLVAXXALDTPQXKKEVQLRKI 181
Db 123 ESEGVVXSXXXIIEHVESVEPLPYXIIDPOLFTKGASLVAXXALDTPQXKKEVQLRKI 182
```

```
QY 182 EXIAQXXXNDYXYITAKEEYKVMNDVARGIVKADVAOSSYGLYQGGQIVAAVDTGLDTG 241
Db 183 EQIAQFALISNDVLYITAKEEYKVMNDVARGIVKADVAOSSYGLYQGGQIVAAVDTGLDTG 242
QY 242 RNDSSMHEAFRGKITALTALGRTNNANDTNGHGVAGSVLGNGATNGMAPOAALVFO 301
Db 243 RNDSSMHEAFRGKITALTALGRTNNANDTNGHGVAGSVLGNGATNGMAPOAALVFO 302
QY 302 IMDSXGGLGGLPSNLQTLFSAQXSAGARIHTNSMGAAVNGATYTDSSRVNDVYRRNDMTI 361
Db 303 IMDSXGGLGGLPSNLQTLFSAQXSAGARIHTNSMGAAVNGATYTDSSRVNDVYRRNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 421
Db 363 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAOFSSRGPKDGR 422
QY 422 KPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVVAQLREHFVNR 481
Db 423 KPDVAPGTIILSARSSLAPDSSFANHDSKRAYMGTSMAPIYAGVVAQLREHFVNR 482
QY 482 GTPPSPSLKKAALJAGADGXGLGYPNGNGMGRTVLDKSLNVAAYNESSXLSSTQKATY 541
Db 483 GTPPSPSLKKAALJAGADGXGLGYPNGNGMGRTVLDKSLNVAAYNESSXLSSTQKATY 542
QY 542 FTATGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXPXKXNMWDGR 601
Db 543 FTATGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYVGNDFXPXKXNMWDGR 602
QY 602 NNVENVFINXPOSGTYTTEVOAYNVPGPQFSLAIYN 639
Db 603 NNVENVFINXPOSGTYTTEVOAYNVPGPQFSLAIYN 640
```

RESULT 3

```
US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
```

```
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-8
```

```
Query Match          97.4%; Score 3021; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.1e-253;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 2 RKKKVFSLVSAALSTVALANPSAGXARXFDLDFKGIQTITDXXGFSKQXOTGAALFL 61
Db 3 KKKKVFSLVSAALSTVALANPSAGXARXFDLDFKGIQTITDXXGFSKQXOTGAALFL 62
QY 62 VESENVKLLKGLKKLETVPANNNKLHIYQNGPILLETQKXILEXTGAKILDIYIPYAYIV 121
Db 63 VESENVKLLKGLKKLETVPANNNKLHIYQNGPILLETQKXILEXTGAKILDIYIPYAYIV 122
```

122	EXEADVXXXXXXIVTSVEPTLPYXXIDPOLFTGAGXKAAALDTRKXNEVOGR1	181	OTHER INFORMATION: xaa	is any amino acid
123	EXEGDVKANTSTIEHVESVEYLPTRIDPOLFTKGASELVKAAALDTRKXNEVOGR1	182	NAME/KEY: misc_feature	LOCATION: (46)..(46)
182	EXIAXKXXSNVXYITAKPEKYAMDVARGTYKADVAASSYGLGCOOLYAADVTDLDIG	241	OTHER INFORMATION: xaa	is any amino acid
183	EQIQAFAISNDVLIITAKPEKYAMDVARGTYKADVAASSYGLGCOOLYAADVTDLDIG	242	NAME/KEY: misc_feature	LOCATION: (47)..(47)
242	RNDSSHEAFRCKITLALYALGRTNNANDNGHGTIVAGSVLGNGXTKGMAPQANLVFOS	301	OTHER INFORMATION: xaa	is any amino acid
243	RNDSSHEAFRCKITLALYALGRTNNANDNGHGTIVAGSVLGNGXTKGMAPQANLVFOS	302	NAME/KEY: misc_feature	LOCATION: (53)..(53)
302	IMDSXGGLGGLPSNLQTLFSGAASAGARIHTNSMGAAVNGAYTTDSRNVDYRKNDMIT	361	OTHER INFORMATION: xaa	is any amino acid
303	IMDSGGLGGLGGLPSNLQTLFSGAASAGARIHTNSMGAAVNGAYTTDSRNVDYRKNDMIT	362	NAME/KEY: misc_feature	LOCATION: (70)..(70)
362	LFAAGXENXNGGTISAGCTAKNNITVGATENLRPSFGSYANINHHVQFSRGPTRDGR1	421	OTHER INFORMATION: xaa	is any amino acid
363	LFAAGXENXNGGTISAGCTAKNNITVGATENLRPSFGSYADININHHVQFSRGPTRDGR1	422	NAME/KEY: misc_feature	LOCATION: (74)..(74)
422	KRPVMAPEGXTLTSARSSLPADSSFFMANHDSKAYVMGGSMTPTVACNVAQLEHEHYKRR	481	OTHER INFORMATION: xaa	is any amino acid
423	KRPVMAPEGXTLTSARSSLPADSSFFMANHDSKAYVMGGSMTPTVACNVAQLEHEHYKRR	482	NAME/KEY: misc_feature	LOCATION: (89)..(89)
482	GTPPKPSLLKAAALAGAADGCLGTPNGNOGMRVTLDSKLVNAVTVNESSLLSTSOARITX	541	OTHER INFORMATION: xaa	is any amino acid
483	GTPPKPSLLKAAALAGAADGCLGTPNGNOGMRVTLDSKLVNAVTVNESSLLSTSOARITX	542	NAME/KEY: misc_feature	LOCATION: (105)..(105)
542	FPTATACKPLKLSLWSDAPASTASVTLVNDLDLVTAPNGTXYVGDFFXXPKXXMMDR	601	OTHER INFORMATION: xaa	is any amino acid
543	FPTATACKPLKLSLWSDAPASTASVTLVNDLDLVTAPNGTXYVGDFFXXPKXXMMDR	602	NAME/KEY: misc_feature	LOCATION: (128)..(128)
602	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	639	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (131)..(131)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (132)..(132)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (133)..(133)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (146)..(146)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (148)..(148)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (160)..(160)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (165)..(165)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (172)..(172)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (183)..(183)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (187)..(187)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (188)..(188)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (189)..(189)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (194)..(194)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (286)..(286)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (286)..(286)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (286)..(286)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (286)..(286)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (286)..(286)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (286)..(286)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	NAME/KEY: misc_feature	LOCATION: (286)..(286)
603	NNNENFTXNPOSCTYTEVOAVNVPGQPSLAIIVN	640	OTHER INFORMATION: xaa	is any amino acid
603				

```
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1
```

Query Match 97.3%; Score 3017; DB 4; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.5e-253;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRRKKVFLSVLSAAALITLVALXNPSAGXARXFDLDFKGIOTTTXXGFSKXOTGAANF 60
DB 1 MRRKKVFLSVLSAAALITLVALXNPSAGXARXFDLDFKGIOTTTXXGFSKXOTGAANF 60
QY 61 LVESENVKLXGKLKLETPANNKLIHXOFNGPILEETKOKLEXTGAKIIDYIPDYAYI 120
DB 61 LVESENVKLXGKLKLETPANNKLIHXOFNGPILEETKOKLEXTGAKIIDYIPDYAYI 120
QY 121 VEYEGDVXSSXXXIEHVESVEPYLPXYXIDPOLFTKGASXLVKAYALDTKOKNKEVOLRG 180
DB 121 VEYEGDVXSSXXXIEHVESVEPYLPXYXIDPOLFTKGASXLVKAYALDTKOKNKEVOLRG 180
QY 181 IEXIXQXXXSNDVXYITAKPEKYVNDVARGIVKADVAQSSGLGGOQIYAVADTGLDT 240
DB 181 IEXIXQXXXSNDVXYITAKPEKYVNDVARGIVKADVAQSSGLGGOQIYAVADTGLDT 240
QY 241 GRNDSMHEAFPGKITLALYALGRTNANDTNGHGHVAGSVYNGCXTKKGMAPOANLVQ 300
DB 241 GRNDSMHEAFPGKITLALYALGRTNANDTNGHGHVAGSVYNGCXTKKGMAPOANLVQ 300
QY 301 SIMDSXGLGLPSNLOTLFQOAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMT 360
```

```
DB 301 SIMDSXGGLGGLPSNLOTLFQOAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMT 360
QY 361 ILFAAGNEXPNGGTISAGTAKNAITVGATENLIRPSFGSYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTISAGTAKNAITVGATENLIRPSFGSYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVAPGTYXILSARSSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHPEYKN 480
DB 421 IKPDVAPGTYXILSARSSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHPEYKN 480
QY 481 KGITPKPSLTKALLAGAADXGLGYPNGNGGRVTLKSLNVAAYVNESSXLTSSOKATY 540
DB 481 KGITPKPSLTKALLAGAADXGLGYPNGNGGRVTLKSLNVAAYVNESSXLTSSOKATY 540
QY 541 XFTATAGKPLKISLWSPDAPASTTASVTLVNDLVTAPNCTXYVGNDFXPPXXNNDG 600
DB 541 XFTATAGKPLKISLWSPDAPASTTASVTLVNDLVTAPNCTXYVGNDFXPPXXNNDG 600
QY 601 RNNVENVEFNKPSGTYTIEVQAVNVPVGPQXFLAIYN 639
DB 601 RNNVENVEFNKPSGTYTIEVQAVNVPVGPQXFLAIYN 639
```

```
RESULT 5
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
```

	LOCATION: (.432)..(.433)	is	any	amino acid
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.502)..(502)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.532)..(.532)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.542)..(.542)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.585)..(.585)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.592)..(.592)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.593)..(.593)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.595)..(.595)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.596)..(.596)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.597)..(.597)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.612)..(.612)			
?	OTHER INFORMATION: xaa			
?	NAME/KEY: misc.feature			
?	LOCATION: (.633)..(.633)			
?	OTHER INFORMATION: xaa			
-S-09-509-814A-2		is	any	amino acid

	Query Match Similarity	97.0%	Score 3007	DB 4	Length 640
	Best Local Similarity	100.0%	Pred. No. 1	9e-252	
	Matches 637	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	3	KKKVFVLSVLSAAALITL	SVALKXNPSGAXRXPDP	PKGIQTTPDXKPSFKOXOTGGAARV	62
Db	4	KKKVFVLSVLSAAALITL	SVALKXNPSGAXRXPDP	PKGIQTTPDXKPSFKOXOTGGAARV	63
Qy	63	SEENKIKLKKGKIKKLETPV	PANNKLIHXOPNPIL	LEFKKXLETKGKILIDYIPDAVIVE	122
Db	64	SEENKIKLKKGKIKKLETPV	PANNKLIHXOPNPIL	LEFKKXLETKGKILIDYIPDAVIVE	123
Qy	123	YEGDVYXXXXXIEHVESV	EPYLYXITDPLQFTYKAS	LKYKAALDTKXKXEVOLRTE	182
Db	124	YEGDVYXXXXXIEHVESV	EPYLYXITDPLQFTYKAS	LKYKAALDTKXKXEVOLRTE	183
Qy	183	XIAQXXXXXNDVYITTA	KPEYKKNVDYANGITY	KADYQSSGTLGGCQIYA	VADTGLDPR 242
Db	184	XIAQXXXXXNDVYITTA	KPEYKKNVDYANGITY	KADYQSSGTLGGCQIYA	VADTGLDPR 243
Qy	243	NDSMHEAFRCKITALT	ALGRYNNANDTNGICHT	IVAGSVLNGCYTKKGA	PAOALVFOST 302
Db	244	NDSMHEAFRCKITALT	ALGRYNNANDTNGICHT	IVAGSVLNGCYTKKGA	PAOALVFOST 303
Qy	303	MDSXGGLGGLPSNLQTL	FSQASAGARITHNSMGA	AVNGAYTTDSRVVDY	RKNDMTIL 362
Db	304	MDSXGGLGGLPSNLQTL	FSQASAGARITHNSMGA	AVNGAYTTDSRVVDY	RKNDMTIL 363
Qy	363	FAAGNEXDNGGTISAP	GTAKAALITVGATENLR	PSFGSYADININHAOFSS	RPTNDGRIK 422
Db	364	FAAGNEXDNGGTISAP	GTAKAALITVGATENLR	PSFGSYADININHAOFSS	RPTNDGRIK 423
Qy	423	PVMAAPGTIXLLSANS	LAPDPSFWANHDSTAY	MGCTISMAPIVAGNAVQ	LRERFVNRK 482
Db	424	PVMAAPGTIXLLSANS	LAPDPSFWANHDSTAY	MGCTISMAPIVAGNAVQ	LRERFVNRK 483

QY 483 ITPKPSLKAALAGADADGGLGYPNGNGWRVTLDKSLNVAAYNESSXLSSTQKATYX 542
Db 484 ITPKPSLKAALAGADADGGLGYPNGNGWRVTLDKSLNVAAYNESSXLSSTQKATYX 543
QY 543 TATGKPKLISLWSDAPASTTASTVTLVNDLVTAPNGXYXVNDXXXPKXXMMDCRN 602
Db 544 TATGKPKLISLWSDAPASTTASTVTLVNDLVTAPNGXYXVNDXXXPKXXMMDCRN 603
QY 603 NVENFINXPOSGTYTIEVQAYNVPVGPQXFLATYV 639
Db 604 NVENFINXPOSGTYTIEVQAYNVPVGPQXFLATYV 640

RESULT 6

US-08-873-479-42
; Sequence 42, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873.479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-873-479-42

Query Match 87.8%; Score 2723; DB 2; Length 641;

Best Local Similarity 82.9%; Pred. No. 7.6e-228;

Matches 532; Conservative 40; Mismatches 66; Indels 4; Gaps 2;

QY 1 MRK--KVLPSLISAAALSTVALXNPSACXARXFDLDFKGIQTITDXXGFSKQXOTGA 57
Db 1 MRKSGKRVFLSVVALLSVALSSPSTIGANNFELDFGIETLLEKAKATQKGTGK 60
QY 58 AALPVSSENKVLAKGLKKLTETVANNKLIHQFNGPILEETKQXLEXGAKIIDYIPDY 117
Db 61 ASFLVNSENVKIPKSIQKLEEVPAADNKLIVCFDGPILLEETQOLETKGAKIIDYIPDY 120
QY 118 AYIYEYGDVYKXXIXLIEHVESVEPYLPXYIIDPOLFTKGSXLYKAAALDTKQXNKVEQ 177
Db 121 AYIYEYGDVAAVTAIAHLESVEPYLPKXIDPOLFSRGASELVEVTAALDKKORSKEVR 180

QY 178 LRGIEXIAQXXXSNDVXYITAKPEYKVNVDVARGIVKADVAQSSYGLXGGOIYAAVADTG 237
Db 181 LRGLQIAQVATNNDVLVTTPKPEYEVLDVARGIVKADVAQNNFGLXGGOIYAAVADTG 240
QY 238 LDTGNDSSMHEAPRGKLTALYALGRTNANDTNHGHVAGSYLGNKXTKKGAPOANL 297
Db 241 LDTGNDSSMHEAPRGKLTALYALGRTNANDTNHGHVAGSYLGN-ATKKGAPOANL 299
QY 298 VFQSTMDXGGLGGLPSMLQTLFSSQAXSAGARIHTNSGAUVNGAYTTDSRNVDDYVKN 357
Db 300 VFQSTMDXGGLGGLPSMLQTLFSSQAXSAGARIHTNSGAUVNGAYTTDSRNVDDYVKN 359
QY 358 DMTILFAAGNEXPMNGTISAPCTAKMALTVGATENLRPSFGSYADNINHVAQSSRGPTK 417
Db 360 DMTILFAAGNEXPMNGTISAPCTAKMALTVGATENLRPSFGSYADNINHVAQSSRGPTK 419
QY 418 DGRKIPDVMAFGTYLXLSRSSIAPDSSFWANHDSKYAAMGTSNATPIVAQNVQALRHF 477
Db 420 DGRKIPDVMAFGTYLXLSRSSIAPDSSFWANHDSKYAAMGTSNATPIVAQNVQALRHF 479
QY 478 VKNGITPKPSLKAALAGADADGGLGYPNGNGWRVTLDKSLNVAAYNESSXLSSTQK 537
Db 480 VKNGITPKPSLKAALAGADADGGLGYPNGNGWRVTLDKSLNVAAYNESSXLSSTQK 539
QY 538 ATYXFTATAGKPKLISLWSDAPASTTASTVTLVNDLVTAPNGTXYXVNDXXXPKXXN 597
Db 540 ATYXFTATAGKPKLISLWSDAPASTTASTVTLVNDLVTAPNGTXYXVNDXXXPKXXN 599
QY 598 WDGRNNEVNFINKPOSGTYTIEVQAYNVPVGPQXFLATYV 639
Db 600 WDGRNNEVNFINKPOSGTYTIEVQAYNVPVGPQXFLATYV 641

RESULT 7

US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873.479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-873-479-43

Query Match 78.0%; Score 2417.5; DB 2; Length 635;
Best Local Similarity 73.3%; Pred. No. 2.2e-201;
Matches 467; Conservative 67; Mismatches 98; Indels 5; Gaps 4;

QY 3 KKKVFLSVLSAAALSTVALXNPSSGAXRFXDLDFKIQTTTXXKXFSKXOTGAALV 62
DB 4 KRRVFLSVLSAAALSTVALXNPSSGAXRFXDLDFKIQTTTXXKXFSKXOTGAALV 60
QY 63 ESENVKXKGLKLETPANKLHIQXFNPIEERKXKLETKGKILIDYIPDAVIVE 122
DB 61 DTEINIPKXIGKLEAVOKXELIYQFTPISEERKGLKESLGLYIDYIPDAVIVQ 120
QY 123 YEGDVXKX 182
DB 121 YSG-ATNINISTLVSENVQPEPLKIDPELTKGASQLQAVYLVNKKENMKFTGLD 179
QY 183 XIAQXKX 242
DB 180 EIVQYKX 239
QY 243 NDSMHEAFKGLKLETPANKLHIQXFNPIEERKXKLETKGKILIDYIPDAVIVE 302
DB 240 NDSMHEAFKGLKLETPANKLHIQXFNPIEERKXKLETKGKILIDYIPDAVIVE 298
QY 303 MDSXGGLGGLPSMLQTLFQSOAXSAGARITHTNSMGAANVAGYTTDSRNVDYRKNDMTIL 362
DB 299 MDSXGGLGGLPSMLQTLFQSOAXSAGARITHTNSMGAANVAGYTTDSRNVDYRKNDMTIL 358
QY 363 FAAGNEXPGNGTISAPGTAKNAITVGATENLRPSFGSYADININHAOFSSRGPTDGRITK 422
DB 359 FAAGNEXPGNGTISAPGTAKNAITVGATENLRPSFGSIADININHAOFSSRGPTDGRITK 418
QY 423 PDVNAPEGTITLSRSLADSSFMANHSKYAVMGSGSMATPIVAGNVQLOLEHFKVARG 482
DB 419 PDVNAPEGTITLSRSLADSSFMANHSKYAVMGSGSMATPIVAGNVQLOLEHFKVARG 478
QY 483 ITPPSSLKALAGAAGXGLGTPNGNOGWRVTLDKSLNVAIVNESSXKLSSTQKATYXF 542
DB 479 ITPPSSLKALAGAAGXGLGTPNGNOGWRVTLDKSLNVAIVNESSXKLSSTQKATYXF 538
QY 543 TATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGITXVYVNDPFXKXKXNDGRN 602
DB 539 TATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGITXVYVNDPFXKXKXNDGRN 598
QY 603 NVENVFINXPOSGTYTIEVQAVNVPCQXESLAIYN 639
DB 599 NVENVFINXPOSGTYTIEVQAVNVPCQXESLAIYN 635

RESULT 8

US-09-104-623A-4
Sequence 4, Application US/09104623A
Patent No. 6303752

GENERAL INFORMATION:

APPLICANT: Olsen, Arne Agerlin
APPLICANT: Falum, Tine Muxoll
APPLICANT: Deussen, Heinz-Josef

APPLICANT: Roggen, Erwin Ludo
TITLE OF INVENTION: A Modified Polypeptide
NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6303752 No. 6303752disk of No. 6303752th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York

STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/104,623A

FILING DATE: 25-JUN-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rozek, Carol

REGISTRATION NUMBER: 36,993

REFERENCE/DOCKET NUMBER: 5256,200-US

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ. ID NO.:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

STRAIN: Bacillus sp. Y

US-09-104-623A-4

Query Match 62.6%; Score 1940.5; DB 4; Length 433;
Best Local Similarity 86.2%; Pred. No. 2.9e-160;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 206 NDVARGIVKADVAKSSGYLGQGOIVAVADTGLDTGRNDSMHEAFKGLKLETPANKLHIQXFNPIEERKXKLETKGKILIDYIPDAVIVE 265
DB 1 NDVARGIVKADVAKSSGYLGQGOIVAVADTGLDTGRNDSMHEAFKGLKLETPANKLHIQXFNPIEERKXKLETKGKILIDYIPDAVIVE 60
QY 266 NANQTNCHGVHAGSVYLGNGXKX 325
DB 61 NASPNCHGVHAGSVYLGNGXKX 119
QY 326 AGARITHTNSMGAANVAGYTTDSRNVDYRKNDMTILFAAGNEXPGNGTISAPGTAKNAIT 385
DB 120 AGARITHTNSMGAANVAGYTTDSRNVDYRKNDMTILFAAGNEXPGNGTISAPGTAKNAIT 179
QY 386 TVGATENLRPSFGSYADININHAOFSSRGPTDGRITKPDVNAPEGTITLSRSLADSSFMANHSKYAVMGSGSMATPIVAGNVQLOLEHFKVARG 445
DB 180 TVGATENLRPSFGSYADININHAOFSSRGPTDGRITKPDVNAPEGTITLSRSLADSSFMANHSKYAVMGSGSMATPIVAGNVQLOLEHFKVARG 239
QY 446 WANIDSRYAAMGTSMAPIVAGNVQLOLEHFKVARGITPKPSLKAALAGAAGXGLGTPNGNOGWRVTLDKSLNVAIVNESSXKLSSTQKATYXF 505
DB 240 WANIDSRYAAMGTSMAPIVAGNVQLOLEHFKVARGITPKPSLKAALAGAAGXGLGTPNGNOGWRVTLDKSLNVAIVNESSXKLSSTQKATYXF 299
QY 506 PNGNOGWRVTLDKSLNVAIVNESSXKLSSTQKATYXFATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGITXVYVNDPFXKXKXNDGRN 565
DB 300 PNGNOGWRVTLDKSLNVAIVNESSXKLSSTQKATYXFATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGITXVYVNDPFXKXKXNDGRN 359
QY 566 SVTLVNDLVLITAPNGITXVYVNDPFXKXKXNDGRNNEVFINXPOSGTYTIEVQAVNVPCQXESLAIYN 625
DB 360 SVTLVNDLVLITAPNGITXVYVNDPFXKXKXNDGRNNEVFINXPOSGTYTIEVQAVNVPCQXESLAIYN 419
QY 626 VPVGPQXESLAIYN 639
DB 420 VPVGPQXESLAIYN 633

RESULT 9

US-09-019-532-4
Sequence 4, Application US/09019532B
Patent No. 6416756

GENERAL INFORMATION:

APPLICANT: Olsen, Arne Agerlin
APPLICANT: Prent, Annette
TITLE OF INVENTION: A Modified Enzyme for Skin Care

FILE REFERENCE: 4922,204-US
CURRENT APPLICATION NUMBER: US/09/019,532B

EARLIER FILING DATE: 1998-02-05
EARLIER APPLICATION NUMBER: 0038/97

EARLIER FILING DATE: 1997-01-10

```

; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 4
; LENGTH: 433
; TYPE: prt
; ORGANISM: Bacillus sp.
; US-09-019-532-4

```

```

Query Match          62.6% Score 1940.5; DB 4; Length 433;
Best Local Similarity 86.2%; Pred. No. 2.9e-160;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

```

```

QY 206 NDVARGIVKADVAQSSYGLGQGIYAVADTGIDTGRNDSMHEAFRGKITLALYALGRIN 265
DB 1 NDVARGIVKADVAQNNYGLGQGIYAVADTGIDTGRNDSMHEAFRGKITLALYALGRIN 60
QY 266 NNDRTGHTHYAGSYLGGXTNKGMARQANLYFOSIMSGGLGGLPSNLQTLFESQANS 325
DB 61 NASDPGHGTHVAGSYLGN-ALNKGMARQANLYFOSIMSGGLGGLPSNLQTLFESQAWN 119
QY 326 AGARHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGGTTISAPGTAKNAI 385
DB 120 AGARHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGGTTISAPGTAKNAI 179
QY 386 TVGATENILRPSFGSYADININHAQESSRGPTRKGRIPVMAPTXILSARSLADSSR 445
DB 180 TVGATENILRPSFGSIADININHAQESSRGPTRKGRIPVMAPTXILSARSLADSSR 239
QY 446 NNNHSHKAYMGCTSMATITVAGNAQALREHFKYKNGITPKPSLLKAALAGADGLGY 505
DB 240 NNNHSHKAYMGCTSMATITVAGNAQALREHFKYKNGITPKPSLLKAALAGADGLGY 299
QY 506 PNGNGMGWHTLDKSLNVAVYNESSXLTSTOKATYFTATAGPKLISLWSDAPASTIA 565
DB 300 PNGNGMGWHTLDKSLNVAVYNESSXLTSTOKATYFTATAGPKLISLWSDAPASTIA 359
QY 566 SYTLVNDLDTLTAPNGTYVGNDEXXPYXXNMDCGNNEVFNXPQSGTYTIEVOAYN 625
DB 360 SYTLVNDLDTLTAPNGTYVGNDESYPYDNMDCGNNEVFNXPQSGTYTIEVOAYN 419
QY 626 VPVGPQXFLAIYN 639
DB 420 VPVGPQXFLAIYH 433

```

```

RESULT 10
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-May-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-818B-1

```

```

Query Match          14.6% Score 451.5; DB 4; Length 659;
Best Local Similarity 25.9%; Pred. No. 5.6e-31;
Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;

```

```

QY 69 LKGLKKLETPYRANKKLIHXDFNGPILEETK--QXLEXTGAKILDYTFDIYATYVEED 126
DB 43 LTRPGFKKQYRNMNMDEVPVTFIMFGSYGDRDRAVYVLRIMGAOV-----KYSYKI----- 92
QY 127 VXSXXXIIEHVESVEEYLYLXXYIDPLFTKGASXLVKAXALDPRKQKREVOQLGIEIXAQ 186
DB 93 -----IYAVAVK-----IKARDLLIAGMIDTGYRG-NTRVSGIKFIQE 130
QY 187 XXXSNDVXYITAKPEXXKNDVARGI--VKADVAQSSYGLGQGIYAVADTGIDTGRND 244
DB 131 -----DYKVOVDATSVSQCADPTVMSLGYDSGVVAVLVDGIDAN--- 173
QY 245 SSMHEAFRGKITALY-ALORTNANNDFNGHGHVAGSYLGNCKXTNK---GMAQNALVPO 300
DB 174 ---HPLDKGRVIGWYDAVNGRSTPYDDGHHGVHAGIVAGTSVNSQYTGVAAGAKLVG 230
QY 301 SIM--DSXGLGLGSLPNTQLTFEQAXSAGARI-----HTNSGAAVNGAYTTD 346
DB 231 KYLGADSGSSVSTIIAGVDMVYQNNKDKYIRVYINLSLSSQSSDGDLSQAVNNAMWA- 289
QY 347 SRNVDDYVRKNDMTILFAAGNEXPNGGTTISAPGTAKNAITVGATENILRPSFGSYADININ 406
DB 290 -----GIVVCVAGANGCPNTYTVGSAAASKVITYGA-----VDSNDN 327
QY 407 VAOFSRGPTRKGRITKPDVMAPTXILSARSLADSSRPMANHDSSYAYMGCTSMATITV 466
DB 328 IASFSRGPTRKGRITKPDVMAPTXILSARSLADSSRPMANHDSSYAYMGCTSMATITV 384
QY 467 ACNVADLRHFVKNRGTIP---KPSLLKAA-LIAGAADGLGYPNGNGMGWHTLDKSL- 521
DB 385 SGVGLALID---AHPSTMDKTKTALILEADIYAPREINDIAY-----GAGRVNYYKALK 436
QY 522 --NVAVNESSXLTSTOKATYFTATAGPKLISLWSDAPASTIASVTLVNDLDTLTA 579
DB 437 YDDYAKLFTGTGAVADKSAHTFEDVSGATFVATLWMD-----TGSDDIDLXLVD 486
QY 580 PNG-----TYVGNDEFXXPXXNMDCGNNEVFNXPQSGTYTIEVOAYNVPVGPX 632
DB 487 PNGNEVDYSYTAAYG-----FEKVGYNPTAGTWTVKVYSK---GAAN 527

```

OY 633 FSLAIYN 639
: : :
DB 528 YQVDVVS 534

RESULT 11
US-09-445-472-12
Sequence 12, Application US/09445472
Patent No. 6358726
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Kunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/09/445,472
CURRENT FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 659
TYPE: PRT
ORGANISM: Thermococcus celer
US-09-445-472-12

Query Match 14.6%; Score 451.5; DB 4; Length 659;
Best Local Similarity 25.9%; Pred. No. 5,6e-31;
Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;
OY 69 LKRLKLETPVANKKLIHXPNGPILEETK--QXLEXTGAKILDYIPDIAYIYEBCD 126
DB 43 LTPLEKRVKVRMMNDEVTVMFESYGRDRAVAVLRMLKMAOV-----KTSKI----- 92
OY 127 VXSXXXXIEHVESVEPYLPXYXIDPOLFTKGASXLVAKXALDTKXNKKEVQLGIEIXIAQ 186
DB 93 -----IPAVAK-----IKARDLLLAGMIDIGYRG-WTRVSGIKFIOE 130
OY 187 XXXSNDVXYITAKPEYKVNDAVARGI--VKADVAOSSYGLYGOGQIVAAVADTGLDTRND 244
DB 131 -----DYKVVDDATSVSOIGADTWNLSIGYDGSQVVVAIVDGDIDAN--- 173
OY 245 SSMHEARCKITALY-ALGRTNNANDTNGHGTVAAGSVLNGXKTNK---GMAPOANLVFQ 300
DB 174 ---HPDLKGVIGWYDAVAVRGSTPYDDGHTVAGIAGTGSVNSOYIGVAPGAKLVGV 230
OY 301 SIM--DSXGGLGGLPNSLQTLFSSQAXSAGART-----HTNSGCAVAVNGAYTTD 346
DB 231 KVLGADGSGSVSTIIAGVDVAVVQNKDKYGIIVNLISGSSOSSDGTDSLQAVNNAMWA- 289
OY 347 SRNVDDVVRKNDMTLFAAGNEXPNNGTISAAGTAKNAITVGATENIRPSFGSYADINH 406
DB 290 -----GLVGVAAAGNSGPRITYVGSPPAAASKYITVGA-----VDSNDN 327
OY 407 VAOFSSNGPKRDRITKPDYMAFGTYIISARBSLAPDSSFMANHDSKATYMGTSNATPIV 466
DB 328 IASSSSGEPFADRLKPEVAVAGVDIAPRAS--GTSKGTPIVDYTTASGTSMTPIV 384
OY 467 AGNVAQLREHFVKNRGITP---KPSLAKA-LIAGAADXGLGYPNGNGQGRVYTLDKSL- 521
DB 385 SGVQALTLQ---AHPSWPTDKAVKKTALLETADIIVAPKEIADIAY---GAGRNVYKAK 436
OY 522 --NVAVNESSKLSQKATYXTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITA 579
DB 437 YDDYAKLTLFGSVAADGSAATHFEDVSGATFYATLTYMD-----TGSSPIDLYITD 486
OY 580 PNG-----TXVYVNDFFXXPXXNMNDRNNVENVFINXPOSGTYITLVEQAVVWPGPOX 632
DB 487 PNGNEVDVSYTAYYG-----FEKVGYNPTAGTWKVKVSYK---GAAN 527

OY 633 FSLAIYN 639
: : :
DB 528 YQVDVVS 534

RESULT 12
US-08-894-818B-35
Sequence 35, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Kunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-894-818B-35

Query Match 13.2%; Score 408; DB 4; Length 654;
Best Local Similarity 25.9%; Pred. No. 3,3e-27;
Matches 162; Conservative 80; Mismatches 210; Indels 174; Gaps 26;
OY 62 VESENVKIKKXKLETPVANKKLIH--IXQFNGPILEETKQXLEXTGAKILDYIPDIAY 119
DB 35 VEKNYGLTLPGLFPRKIOKLPNBEISTVTFPNHREKRIAVRVIELMGARV-----R 87
OY 120 IVEYEGDVSXXXXIEHVESVEPYLPXYXIDPOLFTKGASXLVAKXALDTKXNKKEVQLR 179
DB 88 V-----YHIIIPAI---AADLKVRDLVIGLIGKGRAKLS 118
OY 180 GEXIAQXXXSNDVXYITAKPEYKVNDAVARGIKADVAOSSYGLYGOGQIVAAVADTGLD 239
DB 119 GVRFTQEDYK-----VTVALEGLGDESAQVAVMAYVMNIGYD--GSGITIGITDITGID 170

[illegible]

RESULT 13
 US-09-445-472-16
 Sequence 16, Application US/09445472
 Patent No. 6358726
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, MIO
 APPLICANT: SHIMOTO, Tomoko
 APPLICANT: ASADA, Kiyozo
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 FILE REFERENCE: TAKAKURA-6
 CURRENT APPLICATION NUMBER: US/09/445,472
 CURRENT FILING DATE: 1999-12-06
 PRIOR APPLICATION NUMBER: 151969/1997
 PRIOR FILING DATE: 1997-06-10
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 16
 LENGTH: 654
 TYPE: PRT
 ORGANISM: Pyrococcus furiosus
 US-09-445-472-16

Query Match	13.2%	Score 408;	DB 4;	Length 654;
Best Local Similarity	25.9%;	Pred. No. 3.3e-27;		
Matches 162; Conservative	80;	Mismatches 210;	Indels 174;	Gaps 26

```
0Y      62 VESENVKJXKKLJKLEIVPANNKH--IKQFNPILEETKÖKLEHXGATIDYIDYAV 119
        ||      ||      ||      ||      ||      ||      ||      ||
Db     35 VEKNKGITLPGLFKIKLNPNEISFVIVEENREKEIAVRLELMGARV-----RX 87
0Y      120 IVEYEGDVXSXXXXIEHVSEVPFLPYXXIDPOLFTGASXLKAKALDTRÖKNKEYOXR 179
        |      |      |      |      |      |      |      |
Db     88 V-----YHIIIPAI--AALKRVBDLLIVISGLTGRAKLIS 118
0Y      180 GIEIXIAOXXXSNDXYITAKPEYKVMDVARGIYKADVAQSGLVGQGQIVAVAPTDGLD 239
        | :      | :      | :      | :      | :      | :      | :
Db     119 GVRIIOEDYK-----VTVASLETEGDDESAQAQWARYVMNLGYD--GSGITIIDTGDID 170
```

QY	240	TGRDNSSMHEAFRCKITALTALYLGSTNNAN-----DTNGHGTIVAGSVLGNKXTN-----K	289
Db	171	-----ASHPBLQKV-----IGWVDYVNGSRSTPYDDHGHGHTHVASIAAGTGAASNGKTK	219
QY	290	GMAPOANLVFOSIM--DSXGGLAGLPSNLQTLFQOAXSAGARIHTNSWGA-----	357
Db	220	GMAPKARLAGIKVLGADSGSISTIKIGVEAWVNKKRQGIKVLINLSTSSQSDGTDAL	279
QY	338	--AVNGATTTDSRNVDDYVRKNDMTILFPAAGNEKPNCGTISAPCTAKNATTVGATENLRP	395
Db	280	SOAVNAAMDA-----GLVVVAAGNSGPNKTYTIGSPAASKVLTVGA-----	321
QY	336	SFGSYADNININVAOFSSRGKPKDRIKIPDWAPETXLISARSLAPDSPSFMAHDSXYAY	455
Db	332	-----YDKYDVTITSSSGPLPADRLRLEPVAPENMTIARAS---GTSNGQPIINDYYTA	373
QY	456	MCGTSMATPIYAGVAVQILREHFVYNNRGITRK--PSLKAALLAGA-----ADXGLGT	505
Db	374	APGTSMAPIHVAAGIAALLQ-----AHPSPDPDKVKTALLETADIVRKDEIAD--IAY	424
QY	506	PNGNGMGRVTLDSILNAVYNESSXLSTSOKA-----TYAXPTATAGKPLKISLWSDAP	560
Db	425	-----GACRVNAYRAIN--YDNVAKLVFTGVAKKSGQTHQFIYSGASEVTATLWMDNAN	477
QY	561	ASTPASVTLVNDLDELVTIAPNG-----IYVGNDEXXEPXXXXNMGDNNVENVFNIXPO	613
Db	478	-----SDLDLYLDPDNGNVQVSYTAYYG-----FEKVGYYNPT	511
QY	614	SGTITIIEVOATNVFVGPOXESLAIYN	639
Db	512	DGTWTIKVSYSS--GSANVOVDVNS	534

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: PatentIn Release #1.0, Version #1.30
3
4 CURRENT APPLICATION DATA:
5 APPLICATION NUMBER: US/08/894,818B
6 FILING DATE: 20-MAY-1998
7 CLASSIFICATION: 435
8
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: PCT/JP96/03253
11 FILING DATE: 07-NOV-1996
12
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: JP 323285/1995
15 FILING DATE: 12-DEC-1995
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Broadway, Roger L.
18 REGISTRATION NUMBER: 25,618
19

```

```
; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-818B-5
```

```
Query Match      12.8%; Score 398; DB 4; Length 659;
Best Local Similarity 24.4%; Pred. No. 2.4e-26;
Matches 151; Conservative 94; Mismatches 215; Indels 158; Gaps 24;
```

```
OY 62 VESNNVYKLGKGLKLETPANNKLH--IXQFNGPILEETKQXLETKALIDYIPDAY 119
DB 35 VEKNYGLLPGLERKIQKLNPEEISTVYFENHREKEIAVRLELMGAKV-----RY 87
OY 120 IVEYEGDVYKXXXIIEHVESVEPYLPXYXIDPOLFTKGASLYKAKALDPTQXNKVEQLR 179
DB 88 V-----YHIIPIA--ADLKRDLVLSGLTGKAKLS 118
OY 180 GIEIXAQXXXNDVXYITAKPEYKMDVARGIVKADVAQSSYGLYGQQLVAADTGLD 239
DB 119 GVRFIOEDYK-----VTVSALEGLDESAAQVATYVNMIGYD--GSGITIGIDTGD 170
OY 240 TGRNDSMHEAFRGKITLALYALGRTNNAN-----DTNGHGHVAGSVLNGXTN---K 289
DB 171 -----ASHPDLOGKV-----IGWVDFVNGRSYPYDDHGHGHVSAAGTGAASNGKXK 219
OY 280 GMAPQANLVFQSIM--DSXGGLGGLPSNLQTLFSQAXSAGAR-----HTNSW 335
DB 220 GMAPGALAGIRKVLGDSGSIITIKGVAVDNKDKGIVKINLSGSSOSSDGTDSL 279
OY 336 GAVNGAYITTDSDNNDDYVRKNDMTILEFAAGNEXPGNCTISAPGTAKNAITVATENLRP 395
DB 280 SQAVNNNMDA-----GLVVCVAAGNSGPNYTVGSFAASKRYTGA----- 321
OY 396 SFSGYADNINHAQFSSRGPTKDGRIKPDVMAPTGYILSARSLADSSFMANHDSKYAY 455
DB 322 -----VDSNDNIASFSSRGPTADGRLKREYVAPGVLDLAPRAS--GTSWGPINDYTK 373
OY 456 MGTSMATPIVAGNVAQLREHFVKNRGITP--KPSILKAA--LIAGADXGLGYPNGNOG 511
DB 374 ASGTSMATPIVSGVGCALIQ--APSWTPDKYKTALILETADVAPEKEDIAY-----G 425
OY 512 WGRVYTLDKSL--NVAAYNESXSLSTSOKATYXFTATAGKPLKISLWSDAPASTASVT 568
DB 426 AGRVNVYKAIKDYDAKLTFTGSVADKGSANTHFEDVSGATFVJATLYMD-----T 475
OY 569 LVNDLDLVITAPNG-----TXVYGNDFKXPPXXNMNDGRNENVEFINXPOSQGTTEV 621
DB 476 GSSDIDLVLDPNGNEVDYSTATYTG-----FEKVGYYNPATGATWIVKV 519
OY 622 QAYNVPGQXFSIATVN 639
DB 520 VSTK--GAANTQVDVVS 534
```

```
RESULT 15
US-09-445-472-1
; Sequence 1, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshi
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
```

```
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; US-09-445-472-1
```

```
Query Match      12.6%; Score 391; DB 4; Length 412;
Best Local Similarity 28.9%; Pred. No. 5.2e-26;
Matches 134; Conservative 56; Mismatches 145; Indels 128; Gaps 20;
```

```
OY 223 GLYGQQLVAADTGLDTRNDSMHEAFRGKITLALYALGRTNNAN-----DTNGHGH 276
DB 22 GYDGSQITIGIDTGD-----ASHPDLOGKV-----IGWVDFVNGRSYPYDDHGHGH 70
OY 277 VAGSVLNGXTN--KQMAPQANLVFQSIM--DSXGGLGGLPSNLQTLFSQAXSAGAR 330
DB 71 VASIAAGTGAASNGKYGKMAPGAKLAGIRKVLGDSGSIITIKGVAVDNKDKYIGIRV 130
OY 331 HTNSMGA-----AVNGAYTTDSRNDYVRKNDMTILEFAAGNEXPGNCTISAP 378
DB 131 INLSGSSOSSDGTDLALSQAVNAMD-----GLVVCVAAGNSGPNKTYTIGSP 178
OY 379 GTPAKNAITVATENLRPFSYADNINHAQFSSRGPTKDGRIKPDVMAPTGYILSARSS 438
DB 179 AASKRYTGA-----YDKYDVITSFSSRGPTADGRLKREYVAPGVNMIIAARAS 227
OY 439 LAPDSFMANHDSKAIWGTSMATPIVAGNVAQLREHFVKNRGITP--PILKALIA 496
DB 228 --GTSWGPINDYTTAAPGTSMATPIVAGIALLLQ-----AHPSWTPDKYKTALIE 277
OY 497 GA-----ADXGLGYPNGQNGRVTLDKSLNVAAYNESXSLSTSOKA-----TYXFT 543
DB 278 TADIVKPELID--IAY-----GAGRVAAYKAIN--YDYAKLVFTGVNKKSGQTHQEV 328
OY 544 ATAGKPLKISLWSDAPASTASVTLVNDLDLVITAPNG-----TXVYGNDFKXPPXX 596
DB 329 ISGASFVATILWDMAN-----SDLVLVYDPNGNOVDYSTATYTG----- 369
OY 597 NMDGRNENVEFINXPOSQGTTEVQAYNVPGQXFSIATVN 639
DB 370 -----FEKVGYYNPATGATWITIKVVSYS--GSANYQVDVVS 402
```

```
Search completed: April 1, 2003, 12:42:09
Job time : 17.9891 secs
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:41:04 ; Search time 23.9812 Seconds

(without alignments)
1629.021 Million cell updates/sec

Title: US-09-920-954-1

Perfect score: 3101
Sequence: 1 MKKKVFLSVLSAAAILSTV.....EVOAYNPVGFQMFSLAIVN 639Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*

```
1: /cgn2_6/ptodata/1/pubppa/US08_NEM_PUB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PC7_NEM_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEM_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEM_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_NEM_PUB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US10_NEM_PUB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US60_NEM_PUB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2155	69.5	434	9	US-09-985-689A-1
2	2155	69.5	434	9	US-09-985-689A-2
3	2082	67.1	434	9	US-09-985-689A-6
4	2060.5	66.4	433	9	US-09-985-689A-7
5	1952.5	63.0	433	9	US-09-985-689A-5
6	1948.5	62.8	433	9	US-09-985-689A-3
7	1941.5	62.6	433	9	US-09-985-689A-4
8	451.5	14.6	659	12	US-10-090-624-12
9	408	13.2	654	12	US-10-090-624-16
10	391	12.6	522	12	US-10-090-624-1
11	391	12.6	522	12	US-10-090-624-1
12	307.5	9.9	418	10	US-09-965-921A-2
13	289.5	9.7	1398	12	US-10-090-624-6
14	280.5	9.0	595	9	US-09-927-827-59
15	273.5	8.8	580	9	US-09-927-827-55
16	271.5	8.8	379	9	US-09-813-408-6
17	270.5	8.7	397	9	US-09-779-334A-5
18	255	8.2	382	12	US-10-090-624-31
19	250.5	8.1	380	9	US-09-813-408-16

20	250.5	8.1	380	9	US-09-824-893A-261	Sequence 261, App
21	250	8.1	377	9	US-09-813-408-1	Sequence 1, Appl
22	250	8.1	379	9	US-09-813-408-13	Sequence 13, Appl
23	250	8.1	380	9	US-09-813-408-19	Sequence 19, Appl
24	250	8.1	382	9	US-09-813-408-7	Sequence 7, Appl
25	249.5	8.0	379	9	US-09-813-408-10	Sequence 10, Appl
26	246	7.9	371	9	US-09-813-408-2	Sequence 2, Appl
27	244.5	7.9	379	9	US-09-813-408-11	Sequence 11, Appl
28	244.5	7.9	381	10	US-09-920-118-16	Sequence 16, Appl
29	242.5	7.8	379	10	US-09-920-118-14	Sequence 14, Appl
30	239	7.7	269	9	US-09-976-139-1	Sequence 1, Appl
31	239	7.7	269	9	US-09-976-414-8	Sequence 8, Appl
32	239	7.7	269	10	US-09-837-235-16	Sequence 16, Appl
33	239	7.7	269	10	US-09-860-854B-6	Sequence 6, Appl
34	239	7.7	269	12	US-10-075-907-1	Sequence 1, Appl
35	239	7.7	269	12	US-10-075-895-1	Sequence 1, Appl
36	235	7.6	269	9	US-09-813-408-4	Sequence 4, Appl
37	235	7.6	378	9	US-09-813-408-14	Sequence 14, Appl
38	234.5	7.6	1150	10	US-09-870-122-3	Sequence 3, Appl
39	234.5	7.6	1181	10	US-09-870-122-23	Sequence 23, Appl
40	234	7.5	310	9	US-09-813-408-15	Sequence 15, Appl
41	232	7.5	269	9	US-09-813-408-5	Sequence 5, Appl
42	231.5	7.5	1167	10	US-09-870-122-2	Sequence 2, Appl
43	230.5	7.4	279	9	US-09-813-408-21	Sequence 21, Appl
44	229.5	7.4	372	9	US-09-813-408-9	Sequence 9, Appl
45	229	7.4	269	9	US-09-779-334A-7	Sequence 7, Appl

ALIGNMENTS

```
RESULT 1
US-09-985-689A-1
: Sequence 1, Application US/09985689A
: Publication NO. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGAWA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAKETI, KATSUHIKA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 215483USO
: CURRENT APPLICATION NUMBER: US/09/985.689A
: PRIOR FILING DATE: 2002-07-01
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2001-04-12
: PRIOR APPLICATION NUMBER: JP P2001-114048
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 434
: TYPE: PRT
: ORGANISM: Bacillus sp.
US-09-985-689A-1
Query Match          69.5%  Score 2155:  DB 9:  Length 434:
Best Local Similarity 96.3%:  Pred No. 3 9e-175:
Matches 418:  Conservative 0:  Mismatches 16:  Indels 0:  Gaps 0:
```

QY 206 NDVARKIVADVAQSSYGLYKQGGIVAVADTGIDTRGNSSMHEAFRCKITLALALRTN 255
DB 1 NDVARKIVADVAQSSYGLYKQGGIVAVADTGIDTRGNSSMHEAFRCKITLALALRTN 60
QY 266 NADNTGEGHTVAGSVLNGKXTNKGAPQANLVFQSIMDSXGGLGFLPSNLQTLFSQAXS 325
DB 61 NADNTGEGHTVAGSVLNGKXTNKGAPQANLVFQSIMDSXGGLGFLPSNLQTLFSQAYS 120
QY 326 AGARIFTNSGAANGAVTTDSRVVDYVRKNDMTILFAAGNEXPNGCTISAPCTAKNAI 385

```
Db 121 ACARHTNSWGAAYVAGTTSRNVDDYRKNDMTILFAAGNEGPGGTISAPGTAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINNHVAFSSRGPPTDGRIRKPDVMAPGTYILSARSLADSSF 445
Db 181 TVGATENLRPSFGSYADNINNHVAFSSRGPPTDGRIRKPDVMAPGTYILSARSLADSSF 240
Qy 446 WANHDSKYAYMGSTMAPPIVAGNVAQLREHFVNKNRGITPKPSLLKALIALAGADXLGY 505
Db 241 WANHDSKYAYMGSTMAPPIVAGNVAQLREHFVNKNRGITPKPSLLKALIALAGADXLGY 300
Qy 506 PNGNOGWRVTLDKSLNVAAYVNESSSLSTSQKATYXFTATACKPLKISLWSDAPASTTA 565
Db 301 PNGNOGWRVTLDKSLNVAAYVNESSSLSTSQKATYXFTATACKPLKISLWSDAPASTTA 360
Qy 566 SVTLVNDLDTLTAPNGTYVGNDFXPXXMMWDRNNENVFINKPQGTITIEVOAYN 625
Db 361 SVTLVNDLDTLTAPNGTYVGNDFXPXXMMWDRNNENVFINKPQGTITIEVOAYN 420
Qy 626 VPGPQXFSLATVN 639
Db 421 VPGPQXFSLATVN 434
```

```
RESULT 2
US-09-985-689A-2
: Sequence 2, Application US/09985689A
: Publication No. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGAWA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUOSHI
: APPLICANT: SAKEMI, KATSUHIISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 2154830U
: CURRENT APPLICATION NUMBER: US/09/985, 689A
: PRIOR FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: PRIOR FILING DATE: 2001-04-12
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 2
: LENGTH: 434
: TYPE: PRT
: ORGANISM: Bacillus sp.
US-09-985-689A-2
```

Query Match 69.5%; Score 2155; DB 9; Length 434;

Best Local Similarity 96.3%; Pred. No. 3.9e-175; Indels 0; Gaps 0;

Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy 206 NDVARGIYKADVAOSSYGLYGGOIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 265
Db 1 NDVARGIYKADVAOSSYGLYGGOIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 60
Qy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGGLPSNLQTFLESOAYS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGGLPSNLQTFLESOAYS 120
Qy 326 AGARHTNSWGAAYVAGTTSRNVDDYRKNDMTILFAAGNEGPGGTISAPGTAKNAI 385
Db 121 AGARHTNSWGAAYVAGTTSRNVDDYRKNDMTILFAAGNEGPGGTISAPGTAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINNHVAFSSRGPPTDGRIRKPDVMAPGTYILSARSLADSSF 445
Db 181 TVGATENLRPSFGSYADNINNHVAFSSRGPPTDGRIRKPDVMAPGTYILSARSLADSSF 240
```

```
Qy 446 WANHDSKYAYMGSTMAPPIVAGNVAQLREHFVNKNRGITPKPSLLKALIALAGADXLGY 505
Db 241 WANHDSKYAYMGSTMAPPIVAGNVAQLREHFVNKNRGITPKPSLLKALIALAGADXLGY 300
Qy 506 PNGNOGWRVTLDKSLNVAAYVNESSSLSTSQKATYXFTATACKPLKISLWSDAPASTTA 565
Db 301 PNGNOGWRVTLDKSLNVAAYVNESSSLSTSQKATYXFTATACKPLKISLWSDAPASTTA 360
Qy 566 SVTLVNDLDTLTAPNGTYVGNDFXPXXMMWDRNNENVFINKPQGTITIEVOAYN 625
Db 361 SVTLVNDLDTLTAPNGTYVGNDFXPXXMMWDRNNENVFINKPQGTITIEVOAYN 420
Qy 626 VPGPQXFSLATVN 639
Db 421 VPGPQXFSLATVN 434
```

```
RESULT 3
US-09-985-689A-6
: Sequence 6, Application US/09985689A
: Publication No. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGAWA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUOSHI
: APPLICANT: SAKEMI, KATSUHIISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 2154830U
: CURRENT APPLICATION NUMBER: US/09/985, 689A
: PRIOR FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: PRIOR FILING DATE: 2001-04-12
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 6
: LENGTH: 434
: TYPE: PRT
: ORGANISM: Bacillus sp.
US-09-985-689A-6
```

Query Match 67.1%; Score 2082; DB 9; Length 434;

Best Local Similarity 91.5%; Pred. No. 6.2e-169; Indels 0; Gaps 0;

Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

```
Qy 206 NDVARGIYKADVAOSSYGLYGGOIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 265
Db 1 NDVARGIYKADVAOSSYGLYGGOIYVAVDTGLDGRNDSMHEAFRGKITLALYLGRTN 60
Qy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGGLPSNLQTFLESOAYS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGGLPSNLQTFLESOAYS 120
Qy 326 AGARHTNSWGAAYVAGTTSRNVDDYRKNDMTILFAAGNEGPGGTISAPGTAKNAI 385
Db 121 AGARHTNSWGAAYVAGTTSRNVDDYRKNDMTILFAAGNEGPGGTISAPGTAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINNHVAFSSRGPPTDGRIRKPDVMAPGTYILSARSLADSSF 445
Db 181 TVGATENLRPSFGSYADNINNHVAFSSRGPPTDGRIRKPDVMAPGTYILSARSLADSSF 240
Qy 446 WANHDSKYAYMGSTMAPPIVAGNVAQLREHFVNKNRGITPKPSLLKALIALAGADXLGY 505
Db 241 WANHDSKYAYMGSTMAPPIVAGNVAQLREHFVNKNRGITPKPSLLKALIALAGADXLGY 300
Qy 506 PNGNOGWRVTLDKSLNVAAYVNESSSLSTSQKATYXFTATACKPLKISLWSDAPASTTA 565
Db 301 PNGNOGWRVTLDKSLNVAAYVNESSSLSTSQKATYXFTATACKPLKISLWSDAPASTTA 360
```



```
OY 566 SVTLVNDLDTLTAPNGTXYVGNDFXXPKXXXNMDGRNNVENVFIMXPOSGTYTTEVOAYN 625
D 361 SVTLVNDLDTLTAPNGTXYVGNDFXAPYDNNMDGRNNVENVFIMXPOSGTYTTEVOAYN 420
OY 626 VPGVPOXFSLAIVN 639
D 421 VPGVPOXFSLAIVN 434

RESULT 4
US-09-985-689A-7
: Sequence 7, Application US/09985689A
: Publication No. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGAMA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAEKI, KATSUHIISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 215483050
: CURRENT APPLICATION NUMBER: US/09/985,689A
: PRIOR FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 7
: LENGTH: 433
: TYPE: prt
: ORGANISM: Bacillus sp.
US-09-985-689A-7

Query Match 66.4% Score 2060.5; DB 9; Length 433;
Best Local Similarity 91.7% Pred. No. 4,1e-167;
Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;

OY 206 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 265
D 1 NDVARGIVADVAQNNFGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 60
OY 266 NANDTNGHGTTHVAGSVLGNCKTNKGMAPOANLVFOSIMDSXGGLGRLPSNLQTLFSGQXS 325
D 61 NANDPBGHGTTHVAGSVLGN-ATNKGMAPOANLVFOSIMDSXGGLGRLPSNLQTLFSGQXS 119
OY 326 AGARITHTNSMGAAYNGAVTTDSRNVDYVRKNDMTILFPAAGNEXPNNGTISAPGTAKNAI 385
D 120 AGARITHTNSMGAAYNGAVTTDSRNVDYVRKNDMTILFPAAGNEXPNNGTISAPGTAKNAI 179
OY 386 TYGATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAPGTYILSARSSLAPDSSF 445
D 180 TYGATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAPGTYILSARSSLAPDSSF 239
OY 446 MANHDSKYAYMGSTSMATPIVAGNVADLRHEHFKNKGITPKPSLLKALLAGADXLGY 505
D 240 MANHDSKYAYMGSTSMATPIVAGNVADLRHEHFKNKGITPKPSLLKALLAGADXLGY 299
OY 506 PNGNCGMGRVTLDSKSLNVAAYVNESSXLSTOKATYXFTATGATPKLISLVWSDAPASTTA 565
D 300 PNGNCGMGRVTLDSKSLNVAAYVNESSXLSTOKATYXFTATGATPKLISLVWSDAPASTTA 359
OY 566 SVTLVNDLDTLTAPNGTXYVGNDFXXPKXXXNMDGRNNVENVFIMXPOSGTYTTEVOAYN 625
D 360 SVTLVNDLDTLTAPNGTXYVGNDFXAPYDNNMDGRNNVENVFIMXPOSGTYTTEVOAYN 419
OY 626 VPGVPOXFSLAIVN 639
D 421 VPGVPOXFSLAIVN 434
```

```
DB 420 VPGVPOXFSLAIVH 433

RESULT 5
US-09-985-689A-5
: Sequence 5, Application US/09985689A
: Publication No. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGAMA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAEKI, KATSUHIISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 215483050
: CURRENT APPLICATION NUMBER: US/09/985,689A
: PRIOR FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 5
: LENGTH: 433
: TYPE: prt
: ORGANISM: Bacillus sp.
US-09-985-689A-5

Query Match 63.0% Score 1952.5; DB 9; Length 433;
Best Local Similarity 86.0% Pred. No. 6,1e-156;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

OY 206 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 265
D 1 NDVARGIVADVAQNNFGLYGQGIYAVADTGLDTGRNDSMHEAFRGKITATYALGRTN 60
OY 266 NANDTNGHGTTHVAGSVLGNCKTNKGMAPOANLVFOSIMDSXGGLGRLPSNLQTLFSGQXS 325
D 61 NANDPBGHGTTHVAGSVLGN-ALNKGMAPOANLVFOSIMDSXGGLGRLPSNLQTLFSGQXS 119
OY 326 AGARITHTNSMGAAYNGAVTTDSRNVDYVRKNDMTILFPAAGNEXPNNGTISAPGTAKNAI 385
D 120 AGARITHTNSMGAAYNGAVTTDSRNVDYVRKNDMTILFPAAGNEXPNNGTISAPGTAKNAI 179
OY 386 TYGATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAPGTYILSARSSLAPDSSF 445
D 180 TYGATENLRPSFGSYADNINHVAFQSSRGPTKDGRIKPDVMAPGTYILSARSSLAPDSSF 239
OY 446 MANHDSKYAYMGSTSMATPIVAGNVADLRHEHFKNKGITPKPSLLKALLAGADXLGY 505
D 240 MANHDSKYAYMGSTSMATPIVAGNVADLRHEHFKNKGITPKPSLLKALLAGADXLGY 299
OY 506 PNGNCGMGRVTLDSKSLNVAAYVNESSXLSTOKATYXFTATGATPKLISLVWSDAPASTTA 565
D 300 PNGNCGMGRVTLDSKSLNVAAYVNESSXLSTOKATYXFTATGATPKLISLVWSDAPASTTA 359
OY 566 SVTLVNDLDTLTAPNGTXYVGNDFXXPKXXXNMDGRNNVENVFIMXPOSGTYTTEVOAYN 625
D 360 SVTLVNDLDTLTAPNGTXYVGNDFXAPYDNNMDGRNNVENVFIMXPOSGTYTTEVOAYN 419
OY 626 VPGVPOXFSLAIVN 639
D 420 VPGVPOXFSLAIVH 433

RESULT 6
US-09-985-689A-3
: Sequence 3, Application US/09985689A
: Publication No. US20030022351A1
```

```

: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGWAMA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI
: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAEKI, KATSUHISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 215483050
: CURRENT APPLICATION NUMBER: US/09/985,689A
: CURRENT FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: PRIOR FILING DATE: 2001-04-12
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 433
: TYPE: PRT
: ORGANISM: Bacillus sp.
US-09-985-689A-3

```

```

Query Match          62.8%; Score 1948.5; DB 9; Length 433;
Best Local Similarity 86.4%; Pred. No. 1.3e-157;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

```

```

QY 206 NDVARGIVKADVAOSSYGLYGOGIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 265
DB 1 NDVARGIVKADVAONNNGLYGOGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
QY 266 NANPTNGHGTIVAGSVLGNCGXTNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSAQXS 325
DB 61 NADPNMGHTIVAGSVLGN -ALNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSAQMN 119
QY 326 AGARLHTSMGAAVNGAYTTDSRVNDYVRRNDMTILFAAGNEXPNCGTISAPGTAKNAI 385
DB 120 AGARLHTSMGAPVNGAYTANSRQDEYVRNNDMTVLFPAAGNEGNSGTISAPGTAKNAI 179
QY 386 TVGATENTLRPSFGSYADININVAQFSSRGPTKDGRIKPDVAPGTXILSASSLAPDSSF 445
DB 180 TVGATENTLRPSFGSYADININVAQFSSRGATRDGRIKPDVAPGTXILSASSLAPDSSF 239
QY 446 WANDSKYAYMGSTMATPIYAGNVAOLREHFVKNRGTTTPPSLKAALLIGAADXGLGY 505
DB 240 WANNISKAIYMGSTMATPIYAGNVAOLREHFVKNRGTTTPPSLKAALLIGAADXGLGY 299
QY 506 PNGOGMGRTVLNKSALNVAAYNNESSXLSSTQKATYXFTATAGKPLKISLVNSDAPASTTA 565
DB 300 PNGOGMGRTVLNKSALNVAAYNEATALTGQKATYSFOTQAGKPLKISLVNTDAGSTTA 359
QY 566 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVVFNNXPOSGTYTIEVQAYN 625
DB 360 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVVFNNXPOSGTYTIEVQAYN 419
QY 626 VPVGQFQSLAIYN 639
DB 420 VPVGQFQSLAIYH 433

```

```

RESULT 7
US-09-985-689A-4
: Sequence 4, Application US/09985689A
: Publication No. US20030022351A1
: GENERAL INFORMATION:
: APPLICANT: HATADA, YUJI
: APPLICANT: OGWAMA, AKINORI
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SATO, TSUYOSHI
: APPLICANT: ARAKI, HIROYUKI
: APPLICANT: SUMITOMO, NOBUYUKI

```

```

: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAEKI, KATSUHISA
: TITLE OF INVENTION: Alkaline proteases
: FILE REFERENCE: 215483050
: CURRENT APPLICATION NUMBER: US/09/985,689A
: CURRENT FILING DATE: 2002-07-01
: PRIOR APPLICATION NUMBER: JP P2000-355166
: PRIOR FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: JP P2001-114048
: PRIOR FILING DATE: 2001-04-12
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 433
: TYPE: PRT
: ORGANISM: Bacillus sp.
US-09-985-689A-4

```

```

Query Match          62.6%; Score 1941.5; DB 9; Length 433;
Best Local Similarity 86.2%; Pred. No. 5.3e-157;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

```

```

QY 206 NDVARGIVKADVAOSSYGLYGOGIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 265
DB 1 NDVARGIVKADVAONNNGLYGOGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
QY 266 NANPTNGHGTIVAGSVLGNCGXTNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSAQXS 325
DB 61 NASDPNGHGTIVAGSVLGN -ALNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSAQMN 119
QY 326 AGARLHTSMGAAVNGAYTTDSRVNDYVRRNDMTILFAAGNEXPNCGTISAPGTAKNAI 385
DB 120 AGARLHTSMGAPVNGAYTANSRQDEYVRNNDMTVLFPAAGNEGNSGTISAPGTAKNAI 179
QY 386 TVGATENTLRPSFGSYADININVAQFSSRGPTKDGRIKPDVAPGTXILSASSLAPDSSF 445
DB 180 TVGATENTLRPSFGSYADININVAQFSSRGATRDGRIKPDVAPGTXILSASSLAPDSSF 239
QY 446 WANDSKYAYMGSTMATPIYAGNVAOLREHFVKNRGTTTPPSLKAALLIGAADXGLGY 505
DB 240 WANNISKAIYMGSTMATPIYAGNVAOLREHFVKNRGTTTPPSLKAALLIGAADXGLGY 299
QY 506 PNGOGMGRTVLNKSALNVAAYNNESSXLSSTQKATYXFTATAGKPLKISLVNSDAPASTTA 565
DB 300 PNGOGMGRTVLNKSALNVAAYNEATALTGQKATYSFOTQAGKPLKISLVNTDAGSTTA 359
QY 566 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVVFNNXPOSGTYTIEVQAYN 625
DB 360 SVTLVNDLDVITAPNGXHYVGNDFXAPXXXNMGRNNEVVFNNXPOSGTYTIEVQAYN 419
QY 626 VPVGQFQSLAIYN 639
DB 420 VPVGQFQSLAIYH 433

```

```

RESULT 8
US-10-090-624-12
: Sequence 12, Application US/10090624
: Patent No. US20020132335A1
: GENERAL INFORMATION:
: APPLICANT: TAKAKURA, Hikaru
: APPLICANT: MORISHITA, Mio
: APPLICANT: SHIMOJO, Tomoko
: APPLICANT: ASADA, Kiyozo
: APPLICANT: KATO, Ikuoshin
: TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
: FILE REFERENCE: TAKAKURA-6
: CURRENT APPLICATION NUMBER: US/10/090,624
: CURRENT FILING DATE: 2002-03-06
: PRIOR APPLICATION NUMBER: 09/445,472
: PRIOR FILING DATE: 1999-12-06
: PRIOR APPLICATION NUMBER: 151969/1997
: PRIOR FILING DATE: 1997-06-10

```

NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 12
 LENGTH: 659
 TYPE: PR
 ORGANISM: Thermococcus celer
 US-10-090-624-12

Query Match 14.6%; Score 451.5; DB 12; Length 659;
 Best Local Similarity 25.9%; Pred. No. 3.4e-30;
 Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;

QY 69 LKGLAKKLETVANKKLIHGFNPILEERK--QXLEXTGAILDIYIDYAYIEBED 126
 DB 43 LTPLEFKKQRMNMNOEDVTIMFSSYCDRRAVKVRLMGADV---KYSIKI----- 92
 QY 127 VXSXXXXIHEVESVEPYLPYXIDPDLFTKGSXLYKAKALDTRQXNKKEVLRIGLEXTIAQ 186
 DB 93 -----IPAVAVK---IKARDLLILAGMIDGYEG-NTFVSGIKFIQE 130
 QY 187 XXXSNDVXYITAKPEKKVNDVARGI--VKADVAQSSYGLYGOGQIVAAVADTGLDTRND 244
 DB 131 -----DYKVQVDDATSVQIGADVTWNSLGYDGSQVVAIVDGDIDAN-- 173
 QY 245 SSMHEAFRGKITALY-ALGRTNNANDTNGHGTIVAGSVLNGXTNK---GMAPQANLVEQ 300
 DB 174 ---HPDLKGVIGWDVAVNGRSTPYDDGCHGTIVAGTGSVNSQYIGVAPAKIYGV 230
 QY 301 SIM--DSXGGLGGLPSNLQTLFSSQAXSAGARI-----HTNSGCAVNGCAVTTD 346
 DB 231 KVLGAQSSGSVSTIILGVDMVYQNKDYGIVNLISLSSQSSDGDTSLSQAVNNKMDA- 289
 QY 347 SRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSGYADNINH 406
 DB 290 -----GLVVVAAAGNSGPNVTYVGSPPAASKVITVGA-----VDSNDN 327
 QY 407 VAQSSRGPTKGRIRKPDVMAFPTIILSARSSLAPDSSFPANHDSKYAYMGTSNATPIV 466
 DB 328 IASSSSSGPTPADRLKPEVVAAGVDIIPRAS---GTSMGPTINDIYTTASGSMATPIV 384
 QY 467 AGNVAQLREHFVKNRGIPT--KPSLLKAA-LIAGAAXGLGYPNGNQGMRVTRDKSL- 521
 DB 385 SGVQALTLQ---AHPMTDPKVKYKTLITADIVAPKEIADIAV---GAGRNVYKAK 436
 QY 522 --NVAAYNESSXLSTOKATYFTATAGKPLKISLWSDAPASTFASVTLVNDLDTVITA 579
 DB 437 YDDYAKLITFGSVADKGSATHTFDVSGATFVATLTYND-----TGSSDILYLYD 486
 QY 580 PNG-----TXYVGNDPFXPPXXNMWDRNNVENVFINPQSGTYTIEVOAVNVPRGPX 632
 DB 487 PNGNEVDVSTYATYGG-----FEKVGYYNPTAGTWTYKVSXK---GAAN 527
 QY 633 FSLAIYN 639
 DB 528 XQVDVVS 534

RESULT 9
 US-10-090-624-16
 ; Sequence 16, Application US/10090624
 ; Patent No. US20020132335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, HiKaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 ; FILE REFERENCE: TAKAKURA-6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; PRIOR FILING DATE: 1999-12-06

PRIOR APPLICATION NUMBER: 151969/1997
 PRIOR FILING DATE: 1997-06-10
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 16
 LENGTH: 654
 TYPE: PR
 ORGANISM: Pyrococcus furiosus
 US-10-090-624-16

Query Match 13.2%; Score 408; DB 12; Length 654;
 Best Local Similarity 25.9%; Pred. No. 1.7e-26;
 Matches 162; Conservative 80; Mismatches 210; Indels 174; Gaps 26;

QY 62 VESENVKXKQKKLETVANKKLH--IXQNFPILEETKQXLEXTGAKILDIYIDVAY 119
 DB 35 VEKNVGLTLPGLFPRKIQKLNPEISTVTFEENHEKFIAYVLELGAUV-----RY 87
 QY 120 IVEYEGDVXSXXXXIHEVESVEPYLPYXIDPDLFTKGSXLYKAKALDTRQXNKKEVLR 179
 DB 88 V-----YHILPAI---AADLKVRLDLYISGLTGKRAKLS 118
 QY 180 GIEKIAQXXXSNDVXYITAKPEKKVNDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLD 239
 DB 119 GVRFLQEDYK-----VTVALEGLDESAQVATYVWNLGYD--GSGITIGITDPTGID 170
 QY 240 TGRNDSMHEAFRGKITALYALGRTNNAN-----DTNGHGTIVAGSVLNGXTN-----K 289
 DB 171 -----ASHPDLQKV-----IOWDFVNGRSTPYDDHCHGTIVASIACTGAASNGKK 219
 QY 290 GMAPQANLVEQSIM--DSKGLGGLPSNLQTLFSSQAXSAGARIHTNSGCA- 337
 DB 220 GMAPQAKIAGIKVLAGDGSISTITIKGVNAVNDKDKYIGIKVNLISLSSQSSDGDIDL 279
 QY 338 --AVNGATTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAITVGATENLRP 395
 DB 280 SQAVNAAAMD-----GLVVVAAAGNSGPNVTYVGSPPAASKVITVGA----- 321
 QY 366 SFGSYADNINHVAQSSRGPTKGRIRKPDVMAFPTIILSARSSLAPDSSFPANHDSKYAY 455
 DB 322 -----VKKYDVTISFSSRGPTADGRKPEVVAAGVMIITAAAS---GTSMGQPIYNYA 373
 QY 456 MGSMTAPTIYAGNVQDLREHFVKNRGIPT--PSSLKAAALAGA-----ADXGLCY 505
 DB 374 APGISMATPHVAGIALILO-----AHPSWTPDKVKYKTLITADIVKPDIDIAV---IAY 424
 QY 506 PNGNQGMRVTLDSKLSLVAVNESSXLSTSKA-----TXYFTATAGKPLKISLWSDAP 560
 DB 425 -----GAGRVAAYKALN--YDNVAKLVFTGYVANKGSQTHQFVLSGASFVATLVDNAN 477
 QY 561 ASTASVTLVNDLDTVITAPNG-----TXYVGNDPFXPPXXNMWDRNNVENVFINXPQ 613
 DB 478 -----SDDLTYLPDNGQNDVSTYATYGG-----FEKVGYYNPT 511
 QY 614 SGTITYEVOATYVNGFQXSLAIYN 639
 DB 512 DQWTITKVSYS---GSANYQVDVVS 534

RESULT 10
 US-10-090-624-1
 ; Sequence 1, Application US/10090624
 ; Patent No. US20020132335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, HiKaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 ; FILE REFERENCE: TAKAKURA-6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; PRIOR FILING DATE: 2002-03-06

RESULT 11
 US-10-090-624-4
 Sequence 4, Application US/10090624
 Patent No. US20020132335A1
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: SHIMOJO, Tomoko
 APPLICANT: ASADA, Kiyozo
 APPLICANT: KATO, Ikunoshi
 TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 FILE REFERENCE: TAKAKURA-6
 CURRENT APPLICATION NUMBER: US/10/090,624
 CURRENT FILING DATE: 2002-03-06
 PRIOR APPLICATION NUMBER: 09/445,472
 PRIOR FILING DATE: 1999-12-06
 PRIOR APPLICATION NUMBER: 151969/1997
 PRIOR FILING DATE: 1997-06-10
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 4
 LENGTH: 522
 TYPE: PRT
 ORGANISM: *Pyrococcus furiosus*

```

223 GLYGGGLVAVADPDLDTGRNDSSMHEAFRCCKITLALYLGRTNNAN-----DINGGHT 276
22 GYDGGGILITIGLITID-----ASHPDLQKV-----IGWDFVNGRSRYPDHGHGT 70
277 VAGSVLGNNGXTN-----KMAPQANLVFQSIM--DSXGGLGGLPMLQTLFQSASAGARI 330
71 VASIAAGTGAASNGKRYKGMAPGAKLAGIKVLGADSGSISITLIKVEAVADNKKQYGIKV 130
331 HTNSGA-----AVNGAYTDSRNVDYVRKNDMTILFAAGNEXPNGGTIAP 378
131 INLSIGSSOSSDGTDLALSOAVNAAMD-----GLVVVVAAGNSGPNKTTICSP 178
379 GKAKNAIVGATENLRPEFGSYADNINHAOFSSRGPTKDGRIKPDWAQGXILSARSS 438
179 AAASVYITVGA-----VDKIDYITFSFGPFLADGRLKEVVAAPGNWIIAANAAS 227
439 LAPDSSFANHDSTKAYVNGTSMATPIVAGNVAQUREHFVKNKGIPTPK--PSLLKALIA 496
228 ---GTSMGQPIINDYITAAPGTSMATPRVAGTIALLLQ-----AHPMTDPKVXTALIE 277
497 GA-----ADXGIGTPNGNOGMRVTLDSLNVAIYNESXLSOSKA-----TYAFT 543
278 TADIVKPEIAD--TAY-----GAGRVNAKRAIN--YDNVAKLVETGVANKGSOHQFV 328
544 ATAGKPLKISLWSPASAPASTASVTLVLDLDTVITAPNG-----TYVGNDKFXPPXXX 596
329 ISGASFVATILYWDNAN-----SDLLIYLIDPKNQNOVDYSTLYAYG----- 369
597 NMDGRNNENVEFINXPOSGTYTIEVOAYNVVPPQXFSALAVN 639
370 -----FEKVGYNMPTDGTWTIKVYSY--GSANYQVDVVS 402

RESULT 12
US-09-966-921A-2
; Sequence 2, Application US/09966921A
; Patent No. US20020103100A1
; GENERAL INFORMATION:
; APPLICANT: Outtrup, Helle
; APPLICANT: Lassen, Soren
; APPLICANT: Pedersen, Poul
; TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides Having Proteolytic Activity
; FILE REFERENCE: 10097,200-US
; CURRENT APPLICATION NUMBER: US/09/966,921A
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-966-921A-2

Query Match 9.9%; Score 307.5; DB 10; Length 418;
Best Local Similarity 28.4%; Pred. No.3.1e-18;
Matches 116; Conservative 57; Mismatches 154; Indels 81; Gaps 20;

139 SVEPLPYXIIDPOLFTGASXLVKAALDFTKQNKREYQLRGIEIXIAOXXXSNDYXYITA 198
49 S1STFQSQYDVAKWMDGCKEGFTTDDVAKQLQTLQSNKDIQKVNEMVEYVTTKEAEVTA 108
199 KPEYK-----VMNDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSSMHEAF 251

```

```

Db 109 VPSOTQPMGKISYND--QSITKT-----TGGSGIAVALDHYVTS-----HLLD 152
Qy 252 RGTALVALGRTN-----NANDNGHGTAVAGSVLGNXTN---KQMAQOANI-VFOS 301
Db 153 AGAEQECRKFPTOSNPLVDGCTDRQGHGTHVACTYLAHSGSNGQGVYVAPQAKLMAIKV 212
Qy 302 IMDSXGGLG-GLPSNIQTLFQOAXSAGARITHN-SWGAIVNCAVTTDSRNVDDYVRKNDM 359
Db 213 LGNNGSGYDDIAAARIHVADASRTGSKYVINMISGSSAKSLA---SANDYAYGKV 269
Qy 360 TILFAGNEXPKGRTISAPGTAKNAITVGATENLRSFGSYADNINHYAOSSRG-PTKD 418
Db 270 LIYAAAGNSGSSNTIGFPGGLVNAVAALENVQON-GTY-----RVADFSKGNPATA 323
Qy 419 G-----RIPDYMAPGTYYLSARSSLAPOSSFWANHDSKYAYMGTSNATPTVAGNVA-- 471
Db 324 GDYIIQERDIEVASPGASV-----ESTW--YTGGYNTISGTSNATPTVAGLAAKT 371
Qy 472 -----QLREHFVNKRGITRPSLLAKALIAGAAD--XGLGYP 506
Db 372 WSANTSLSHSLRTE-LQNR--KYVDIKGIGAGTGDGDVAGSGFYP 415

```

RESULT 13

```

US-10-090-624-6
; Sequence 6, Application US/10090624
; Patent No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHTA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshi
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1398
; TYPE: PRN
; ORGANISM: Pyrococcus furiosus
US-10-090-624-6

```

```

Query Match          9.7%: Score 299.5; DB 12; Length 1398;
Best Local Similarity 22.9%: Pred. No. 7 6e-17;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;

```

```

Qy 96 LEETKQXLEXTGAKILD-----YIPDYATIVERE-----GDVXSXXXXIE 135
Db 77 LBAKTELEKLAETLIDENRYLNLMLVKIRPEYVKELNTISLEKAMLINREVKLSPTIVE 136
Qy 136 -HVESVEPLPYXIXIDPOLFT-----KGA SXLVKAXALDTOKXNKEV 176
Db 137 KDKVTEP-----SLEPMYMTWVINALQTFQEGYDSSGVVAVLDGTVDPMHPELSI 191
Qy 177 QLRGIEIXIAQXXSNDVYXITAKPEY-KVNN-----DVARGI----- 212
Db 192 TPDGRKRIIEEMKDFDEGFVDTSPFSKVVNGTLIINTTFQVAGSLGTLESTGLMEYVK 251
Qy 213 -----VKADVAQSSYGLY-----GQGOVVAADT 236
Db 252 TVYVSNVITIGNITSANGIYHGLLPERYFDLNFQDQEDFYVLLVNSTGCGYDIAYVDT 311
Qy 237 GLDTGKNDG-----SMHEAFRGKITALYALGTTNNAN-----DINGCHGTIVAG 279
Db 312 DLQYDPTDEVPYLGQYNTYVDVAVFSYYTGPLNLYVLAETIDPNGEYAVPFGMDHGHTIVAG 371

```

```

Qy 280 SVLNGNXTN-----KGNAPQANLVPQIMDSXG 307
Db 372 TVAGYDSNNDADWMLMSYSGEMEVEFSRLYGMVDYNTVTDVQVAPQAIMAIRVLS-D 430
Qy 308 GAGGLPSNIQTLFQOAXSAGARITHNSWGAIVNCAVTTDSRNVDDYVRKNDM 359
Db 431 GSGSMMDIIECK-YYAANTHGAADYJMSLGG--NAPYLDGTIPRESVAADLEKRYGVAFYI 487
Qy 364 AAGNEPKNGRTISAPGTAKNAITVGATENLRSFGSYADNINHYAOSSRG-PTKD 418
Db 488 AAGNEPKNGRTISAPGTAKNAITVGATENLRSFGSYADNINHYAOSSRG-PTKD 418
Qy 406 HAAQSSRQPTKGRIRPDYMAPGTYYLSARSSLAPOSSFWANHDSKYAYMGTSNATPTV 465
Db 546 RIAPFSRQPTKGRIRPDYMAPGTYYLSARSSLAPOSSFWANHDSKYAYMGTSNATPTV 465
Qy 466 VAGNVAQLREHFVNKRGITRPSLLAKALIAGAAD--XGLGYPNGQNGRYTLDK 519
Db 596 VSGVYVALLISG-AKAEGIYYNPDIKRYLESQATWLEGGDPYTGQRYTELDOGHLYNVTK 654
Qy 520 SLNVAVYVNESSXLSSTQKATYXFTATAGRPKLSLWSDAPASTTA--SVTLVNDLIDLV 576
Db 655 SWEI-----LKAINGTTLPIVDHMDKRSYDFAEYLGVDYIRGLYAR 696
Qy 577 ITAPN-----GTYVGN---DPXXPXXXXNN-----DG-----RNVE----- 605
Db 697 NSIPDIEWHIKIYVGDIEKRFELIATEEMIKRPFVSGVILENTEFVLRKYYDVGLEP 756
Qy 606 -----NVFINXPOSGTYTIEVOAYNVPNGRQXS 634
Db 757 GLYVGRHIIIDPT--TPYIEDELINTVILPEKT 788

```

RESULT 14

```

US-09-927-827-59
; Sequence 59, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseler, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 59
; LENGTH: 595
; TYPE: PRN
; ORGANISM: Xanthomonas campestris
US-09-927-827-59

```

```

Query Match          9.08: Score 280.5; DB 9; Length 595;
Best Local Similarity 24.9%: Pred. No. 9.9e-16;
Matches 125; Conservative 64; Mismatches 163; Indels 151; Gaps 21;

```

```

Qy 92 NGPILTEKQXLEXTGAKILDYIPDYATIVEREGDVXSXXXXIEHVESVEPLPYXIXIDP 151
Db 136 NAVILKTLSEDLAEHGA-----DYRHEVLEILD-----RRVIGAP---AGVNA 176
Qy 152 QLETFGASLVKAXALDTOKXNKEVQLRG-----HEIXAQXXSNDVXYIT 197
Db 177 QLTIHNRQAGHSTGLDGDALDEATELEDSLQRLISARLTRRETILOSRYRALGPRRVA 236
Qy 198 AKPEYKVNNDVARGIVKADVAQSSYGLYXGQOIVAVADPTGIDTGRNDSMHKAF-----R 252
Db 237 RDAGKRALIRHSGVNLVNDAAARTAYBARQOIGMAYLDGT-----AAAHPRFVPPGR 290
Qy 253 GKITALY-----ALRTNNAADTNGKGTIVAGSYLG-----NGXTNR- 289
Db 291 DNVAQMDCTRRGAPKRLTRADGRKRFPTALDGHGHTIAGITACGCAAVPADXGKPGEL 350

```

QY 290 ----GMAPQANLVFQSTIMDSXGLGLPS-----NLOTLEFSQA-----XSAGAR 329
DB 351 LEFGAMAPETQLYGFYKVLDDAG--NGRDSWMKAKAVOHADINERAGELIVHGNLSLGGY 408
QY 330 IHTSMGCAAVGATTTDSRANDDVVRKNDMTILFAAGNE-----XPNCGT-----I 375
DB 409 FDPESYCGGF-----TPLCNELRLMRQGVLLVVAAGNEGLAMLQNGGTYPPANMDLSI 463
QY 376 SAPGTAKNAITVGATENLRP-SFGSYADININHAQFSSRCPTKDRIKPDVMAFGTXILS 434
DB 464 SDPQNLEDAIVVGSVHNSPPHNYG-----VGFSSRGFTADGRSKPDVVAAGEKLS 515
QY 435 ARSSLADPSSFWANHDSKYAYMGTSMTAPTYAG-----NVAQL-----473
DB 516 AAYGFDP-----RDPSSILVEMSGTSMAPHPVSGVLGFLSARREIFGPDPRVKQLLDT 570
QY 474 -----REHFVKNRGITTPKPSILK 491
DB 571 CTDLQRDRXYOGRGV--PNIIMR 590

RESULT 15

US-09-927-827-55
; Sequence 55, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseier, Thomas M.
; TITLE OR INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(13824)B
; CURRENT APPLICATION NUMBER: US/09/927, 827
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 55
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-09-927-827-55

Query Match 8.8%; Score 273.5; DB 9; Length 580;

Best Local Similarity 22.7%; Pred. No. 3.8e-15;

Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;

QY 1 MRKKKFLSVLSAAAILSTVALNPSAGXARXFDLDFKIQIOTTTDXGFSKQXOTGAAP 60
DB 6 LRKRTGSLTILGASALTSLLAMPAPAG---EYILDGLATPOTHOKEFIVTYKDGSTAL 60
QY 61 LVESENVKLXKGLXKLETPVANKKLHXQFNCPILTEETKQXLEXTGAKILDYIPDYAYI 120
DB 61 ASPS---ALTSLRTAANAAPA-----KAGKALGINSYRLALGPELV-- 100
QY 121 VEYBGDVXSSXXXIEHVESVEPYLPXYIIDPOLFTKGASXLVKAXALDTKQXNKEVOLRG 180
DB 101 -----RADRALDRAEFTLMROLAADPNVQSYEV----- 129
QY 181 IEXIAQXXXSNDVXYITAKP-EYKVMNDVARGIYKA--DVAQSSGLYGGOIYAVADTG 237
DB 130 -----DQILHAITLTPNDTRLSEQMAFGTTNAGLINIRPAMDKATGSGVAVADTG 179
QY 238 L-----DTGRNDSMHEAFRGKITALYALGRTNAND-----TNGHGT 275
DB 180 ITSHADLNANILAGYDFISDATTARDGNGRDSNMADEGDWYAANCGAGIPRASSSWHGT 239
QY 276 HVAGSV--LGNGXTN-KGMAPQANLVFQSTIMDSXG-----LGLPSEN 316
DB 240 HVAGTVAAVTNNNTGTGAVGAKVVPYVLGKGGSLDAIDAIWVASGGTVSGIIPANA 299
QY 317 Q--TLEFSQASAGARIHTNSMGCAVNGAVYTTDSRANDDVVRKNDMTILFAAGNEXPNGCT 374
DB 300 NPAEYINSLGSGGSCSTTMQN-AINGAVNSRGT-----TVVVAAGNDASNVSG 346

QY 375 ISAPGTAKNAITVGATEN--LRPSFGSYADININHAQFSSRCPTKDRIKPDVMAFGTXI 432
DB 347 -SLPANCANVAIVAAATTSAGAKASYSNFGTGI-----DVSAPGSSI 386
QY 433 LSARSS--LAPDSSFMANHDSKYAYMGTSMTAPTYAGNVAQLEHFVKNRGITPK--PS 488
DB 387 LSTLNSGTTTPGS-----ASVASYNGTSMASPHVAGVALVOS--VAPALTPAAVET 437
QY 489 LLK--AALLAGMADXLGY-----PNGNGWGRVTLDKSLNVAAYNESX 531
DB 438 LKKTATRALPGACGGGAGIVNADAIVTAIINGSGGGGGGNTLNGTPTVTLG---- 493
QY 532 LSTQKATYXFTATAGKPKLTISLWSDAPASTTASVTL--VNDLDVY--TAPNGTXY 585
DB 494 AATGAELNVTITVPAG-----SGTLVTVTSGGSGDADLYVRAGSAPDTSAY 539
QY 586 VGNDPFXPXXNMMDGRNVENVFINXQSGTYTTEVOAYN 625
DB 540 TCRPYRS-----GNAETCTITAP-SGTYVYVRLAYS 569

Search completed: April 1, 2003, 12:52:24
Job time : 26.9812 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:36:24 ; Search time 17.4863 Seconds
(without alignments)
3513.026 Million cell updates/sec

Title: US-09-920-954-1

Perfect score: 3101

Sequence: 1 MRKKVFSLVLSAAAILSTV.....EQQAVNPVGPQXSLATVN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	518.5	16.7	1743	2	T18279	multidrug resist
2	477.5	15.4	1905	2	T18267	intracellular alka
3	350.5	11.3	444	2	B83891	intracellular alka
4	319.5	10.3	442	2	A69587	pyrolysins (EC 3.4.
5	297.5	9.6	1398	2	T28159	microbial serine p
6	286	9.2	806	2	A41341	subtilisin (EC 3.4
7	279.5	9.0	419	1	S25835	subtilisin-type pr
8	274.5	8.9	799	2	G83753	subtilisin (EC 3.4
9	273.5	8.8	420	1	S23407	serine proteinase
10	268.5	8.7	580	2	JC4908	cell wall-associat
11	262	8.4	894	2	F69730	surface layer-asso
12	260.5	8.4	1345	2	T29090	high-alkaline seri
13	257.5	8.3	378	2	A33973	subtilisin (EC 3.4
14	255.5	8.2	381	2	JH0778	subtilisin (EC 3.4
15	253.5	8.2	513	1	A35742	subtilisin (EC 3.4
16	253.5	8.2	381	1	SUBSN	subtilisin (EC 3.4
17	253.5	8.2	381	1	SUBSN	subtilisin (EC 3.4
18	253.5	8.2	381	1	SUBSN	subtilisin (EC 3.4
19	251.5	8.1	381	1	SUBSN	subtilisin (EC 3.4
20	250.5	8.1	380	2	JC4576	high-alkaline seri
21	250.5	8.1	601	2	JC4576	serine proteinase
22	250	8.1	382	2	I39780	subtilisin (EC 3.4
23	249.5	8.0	379	1	SUBSCL	subtilisin (EC 3.4
24	249.5	8.0	757	2	C84120	subtilisin-type pr
25	248	8.0	627	2	D75393	serine proteinase
26	247.5	8.0	381	2	J01487	subtilisin (EC 3.4
27	246.5	7.9	402	2	JU0332	alkaline proteinase
28	242	7.8	1331	2	A72647	probable surface 1
29	239.5	7.7	401	2	I39974	serine proteinase

30	239.5	7.7	534	1	J50173	alkaline proteinase
31	235	7.6	519	2	S71451	halolysin R4 (EC 3
32	232	7.5	488	2	A11930	protease [import
33	231.5	7.5	1167	1	A33066	streptococcal C5a
34	230	7.4	384	2	JC4802	alkaline proteinase
35	223.5	7.2	409	1	S32905	serine proteinase
36	223.5	7.2	321	1	S27501	alkaline proteinase
37	220.5	7.1	535	2	B82358	hypothetical prote
38	219.5	7.1	613	2	S75976	subtilisin (EC 3.4
39	218.5	7.0	272	2	A23624	subtilisin (EC 3.4
40	218.5	7.0	275	2	JC1085	subtilisin-type al
41	218.5	7.0	533	1	JU0146	serine proteinase
42	218	7.0	361	2	G83756	subtilisin-type al
43	216.5	7.0	361	2	A48373	high-alkaline seri
44	216.5	7.0	525	2	G84406	halolysin [importe
45	216.5	7.0	645	1	SUBSMP	serine proteinase

ALIGNMENTS

RESULT 1

T18279

multidrug resistance transport protein - slime mold (Dictyostellium discoideum)

C:Species: Dictyostellium discoideum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18279

R:Shaulsky, G.; Loomis, W.F.

Submitted to the EMBL Data Library, June 1996

A:Reference number: Z18855

A:Accession: T18279

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1743 <SHA>

A:Cross-references: EMBL:U60086; NID:g1399914; PID:g1399915; PIDN:AAB03331.1

C:Genetics:

A:Gene: tagC

Query Match

Best Local Similarity 16.7% Score 518.5 DB 2; Length 1743;

Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

QY	87	HIQFNGPILEETKQXL-----EXTGAKILDYIPDVAYIV--EYEGDYXN 130	
DB	115	YIVQFKRIDNDEFREKLEFLIGTDIVDEQPYOSHIVYIPHDSFLVLMQEOSVLLS 174	
QY	131	XXXIEHVESVEPY-----LPYXXIDPOLFTGASXLVK--AXALDT--KQXN 173	
DB	175	KEVWSVIGFEPSNKHILVNEKSGILPYLI---LSDSTNSLIQREMENTLNSILKSVN 230	
QY	174	KEYOLRG-----EXINQXXXNDVXYXTAKPEVK 203	
DB	231	SKYKLLINQKRLKSYVNCNDESPSPCSLINSERKLYOMISEQSEN--FIRSEKTD 287	
QY	204	VMDVARGIV-----KADVAQSSYGLYGQQLVAVADTGLDTGR---NDS----- 245	
DB	288	TARRLSPKVFETKDTLVNNDRDIP---LKGQGLISTIDTGLDSCHEFSQSYVPI 342	
QY	246	---SMHEAFRGKITVALVGLGTTNANNDTNGHGVAGSYLG-----NGXTNKGMAPOA 295	
DB	343	PLNSVNLNR-KVIVYIYITSTSDSDKVDHGHILGSAAGPREDSSVNTSFSGLATPA 401	
QY	296	NLVFQSIQKSGLEG---PSNLOTFLSQAXSAGARHNTNSGA---AVNGAYTTDSRN 349	
DB	402	KIAF---FDLAGSSSLTPPSDLKOLYOPLYDAGARVHDSGVSVEGYGYSPTAS 458	
QY	350	VDDYVARN-DMTILFAAGNEXNPGTIS--APGTAKNATVVGATENLR-----PSFCS 399	
DB	459	IDDFLTHPDIILRAAGN---NEQYISLTOSTAKNVITVGAGHQTIBENVLDGPNYIN 515	
QY	400	YADNI----- 404	
DB	516	YQSSVDINELICDPDSRYCNYTTAQCCLSNATGCLASCCPTLLKRSVIDANNTQPLLY 575	

Db 318 -RDTTDEDD---VAPSSRGPTTYGKPKFDILAPGVNIVLSRS---PNSFYDKIQKGS 370
Oy 448 NHDSKYAYMGSTMATPIVAGNVNQLREHFVKNGITPKPSLKAALLAGA---AD--- 500
Db 371 RVGSHHTYMGSTSMATPVAGVVALMLQH---EPNLTPOE---VKTRLMGTDRMADRPDN 425
Oy 501 -XGLGY-----PNCNQ 510
Db 426 VYGAGYISAGCAIPNSEE 443

RESULT 4

intracellular alkaline serine proteinase aprx - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: A69587

R:Kunst, R.; Ogatawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertei

A.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Gallizzi, A.; Galler

lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koether, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadachi, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danclon, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044003; PMID:9384377

A:Accession: A69587

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1442 <KUN>

A:Cross-references: GB:299113; GB:AL009126; NID:g2634090; PIDN:CAM13610.1; PID:el183365;

A:Experimental source: strain 168

C:Genetics:

A:Gene: aprx

C:Superfamily: subtilisin homology

F:146-398/Domain: subtilisin homology <SBT>

Query Match 10.3%; Score 319.5; DB 2; Length 442;

Best Local Similarity 30.0%; Pred. No. 4.4e-15;

Matches 106; Conservative 46; Mismatches 108; Indels 93; Gaps 15;

Oy 201 EYKMNDAVAGIYKA-DVAOSSYGLYGQGIYAVADTGLDTGRNDSMEAFRGKITALY 259
Db 120 EVALLDITATEASHAKEVVRNQGTLTGKGYVAVVDIGI-----YHPDLEGRI----- 168
Oy 260 ALGRTNAN-----DINGHGHVAGSVLNGXTN-----KGNAPQANLVFQSIIMDSXGGL 309
Db 169 -IGFADVNVOKTEPYDDNNGHGHGAGDYASGASSGQYGRPAENLGVVLNKKQGS- 226
Oy 310 GGLPSNLQTLFQSAXSAGARITHNSMGAAVY-----GAYTDSRANVD 352
Db 227 GTLADIIEGV-----EMCIQYNEDNPDEPIDIMSKLGGALXYDHEQDP 272
Oy 353 YVKNKD-----MTILFAAGNEXPNCGTISAPGTAKAITYGTEMLRPSFGSYADNINH 406
Db 273 LVRAVEANSAGIVCAVAGNSGPDQTIASPGSEKIVITVGLADNNTA-----SSDDT 328
Oy 407 VAOFSRGPTKDGRIKPDVMAFGTYILSARSLPADSSF-----NANHDSKYAYMGSTSM 461
Db 329 VASSSSGCPVYGKERPDILAGVNIISLRS---PNSYIDKLOKSSRSVQSOFTMGSTSM 385
Oy 462 ATPVAGNVAQLREHFVKNGITPK--PSLKAALLAGAAXGCLPNCNQMG 512
Db 386 ATPICAGIAALILQ---QNPDLTPEVKEVLLK-----NCTDKM 420

RESULT 5

pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus

C:Species: Pyrococcus furiosus

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000

C:Accession: T28159

R:Voornhorst, W.G.B.; Eggen, R.J.L.; Geerling, A.C.M.; Platteuw, C.; Slezén, R.J.; Vo

J. Biol. Chem. 271, 20426-20431, 1996

A:Title: Isolation and characterization of the hyperthermostable serine protease, pyr

A:Reference number: 220481; MUID:96355370; PMID:8702780

A:Accession: T28159

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 11398 <VOO>

A:Cross-references: EMBL:U55835; NID:g1556462; PID:g1556463; PIDN:AA809761.1

A:Experimental source: DSM3638

C:Genetics:

C:Keywords: hydrolase; serine proteinase

Query Match 9.6%; Score 297.5; DB 2; Length 1398;

Best Local Similarity 22.9%; Pred. No. 7.6e-13;

Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;

Oy 96 LEETKQXLEXTAKLID-----YIPYAYIVEY-----GDYXSXXXIE 135
Db 77 LEEKTELEKLAELIDENRYLNLVIRKPEKVELNYSLEKAMLNREKXLSPIVE 136
Oy 136 -HVESVPEPLPYXIDPOLFT-----KGSXLVKAAXALDTQXKKEV 176
Db 137 KDVTKPEP-----SLEPKMYSTWVYINNALQFIQEPFGDSGVVAVVLDTCVDPNHPFLSI 191
Oy 177 QLGIEKIXQXXNSNDVYVYTRAKREV-KVNN-----DVARGI----- 212
Db 192 TPQGRKKIIEKMDFTDEGFVDTSEFSKVVYNGTLLINTTPQASGLTLNESTGLMEYVK 251
Oy 213 -----YKADVAOSSTGLY-----GGQGIYVADT 236
Db 252 TVYVSNTYIGNTISANGYIHFGLLPERFPDLNFGQDEDPYLLVNSTGNGYDAIYDT 311
Oy 237 GLDTGRNDS-----SMHEAFRGKITALYALGRTNAN-----DINGHGHVAG 279
Db 312 DLVYDTDEVPILQYNTVYDVAVSYYGPLNYVLAIEDPNEGYAVFGMDCHGHTHVAG 371
Oy 280 SVLNGXTN-----KGNAPQANLVFQSIIMDSXG 307
Db 372 TVAGYSNNDAMDMLSMYSGMEVFSRLGYMDTYNTDTFVQGVAPGAQIMAIRVLS-D 430
Oy 308 GLGGLPSNLQTLFQSAXSAGARITHNSMGAAVNGAYT--TDSRN--VDDYVRKNDMTILF 363
Db 431 GRGSMMDIIEGM-TYAAVTHGADVISMISLGG--NAPYLDGTDESAVDELTEKYGAVFVI 487
Oy 364 AAGNEXPNCGTISAPGTAKAITYGATEINLRPSFGSYAD-----NIN 405
Db 488 AAGNEXPNCGTISAPGTAKAITYGATEINLRPSFGSYAD-----NIN 405
Oy 406 HVAQFSRGPTKDGRIKPDVMAFGTYILSARSLPADSSFMANHDSKYAYMGSTSMATPI 465
Db 546 RIAFSSRGRIDELKRPVAVAPGIGYISLPMWIGADP-----MSGISMATPH 595
Oy 466 VAGNVAQLREHFVKNGITPKPSLKAALLAGA-----DXGLGYPNCNQMGKRVTLTK 519
Db 596 VSGVYVALLISG--PKRPGIYVNDIIRKVLKESGATWLECGPYTGQKRTLEDGNGIYVNTK 654
Oy 520 SLNVAAYVNESSXLSTSORATYXFTATAGKPLKISLVMSDAPASTTA---SVTLVNDLDLV 576
Db 655 SWEI-----LKAINGTTLPIVDHMAKSYSDAEVILGVDVIRGLYAR 696
Oy 577 ITAPN-----GTYYVGN-----DFXXPXXXNN-----DG-----NNNYE----- 605
Db 697 NSLPIDLVEWHIKRYVGDTEYRTFEIYATEPMIKPFVGSVYLNNMTEFVLRVKYVDEGLEP 756
Oy 606 -----NVFLINXPOSGTYTIEVOAVNPVGPQXFS 634

Db 757 GLYVGRITIDPT--TPVIEDLINTIPIPEKFT 788

RESULT 6

A41341 microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis

C/Species: Bacillus subtilis
C/Date: 17-Jul-1992 #sequence,revision 17-Jul-1992 #text_change 20-Jun-2000
C/Accession: A41341; B41341; S39700; D69730

R/Strom, A.; Rufo Jr, G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.

J. Bacteriol. 173, 6889-6895, 1991

A>Title: Cloning and characterization of the gene for an additional extracellular serine

A/Reference number: A41341; MUID:92041574; PMID:1938892

A/Accession: A41341

A/Molecule type: DNA

A/Residues: 1-806 <SLD>

A/Cross-references: GB:M76590; NID:g143819; PIDN:AAA22881.1; PID:g143820

A/Accession: B41341

A/Molecule type: protein

A/Residues: 161-195 <SL2>

R/Glasier, P.; Kunst, F.; Arnaut, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,

A., Rapoport, G.; Danchin, A.

Mol. Microbiol. 10, 371-384, 1993

A>Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr

A/Reference number: S39655; MUID:95020537; PMID:7934828

A/Accession: S39700

A/Molecule type: DNA

A/Residues: 1-806 <GLA>

A/Cross-references: EMBL:X73124; NID:g413923; PIDN:CA551601.1; PID:g580871

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Buschli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A.; Authors: Lader, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogihara, A.; Oudegaa, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Potetelle

Rieger, M.; Rivolta, C.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A/Authors: Schlecht, S.; Schroter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: D69730

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-806 <KUN>

A/Cross-references: GB:299123; GB:AL009126; NID:g2636240; PIDN:CAB15835.1; PID:g2636344

A/Experimental source: strain 168

A/Comment: The amino terminal sequence of the mature protein and a molecular weight of C

C/Genetics:

A/Gene: vpr

A/Start codon: TTG

C/Superfamily: microbial serine proteinase vpr; subtilisin homology

C/Keywords: hydrolase; serine proteinase

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-160/Domain: propeptide #status predicted <PRO>

F:160-548/Domain: subtilisin homology #status atypical <SBT>

Query Match 9.2%; Score 286; DB 2; Length 806;

Best Local Similarity 22.5%; Pred. No. 2.4e-12;

Matches 132; Conservative 60; Mismatches 186; Indels 206; Gaps 19;

155 TKGASLVKAYALDTKQXNKVEQ--LRGI-----EXIAQXXSNDVXYITAKPEYKVMN 206

Db 87 TKANKKAIK--AVNGKVNREYEOVSGFSMKLPANKEIPKLLAVDVAAYVPMVYTKFDN 144

QY 207 DVANGIVADVAQS-----SYGLYGQGOIYAVVADTGIDTGR----- 242

Db 145 MKDKDVTISDVAVSQMDSDAPYIGANDAWDLGYTGKIKVAILDITGVEYNHPDLKKNFG 204

QY 243 -----NDSSMHEAFRGKITALVALGRTNNANDNGHCTHYAGVGLNGXTKGMAP 293

Db 205 QYKGYDEVNDNDYDPKEPTGT-----DPRGEATDHTGTHVAGTVANG--TIKGVAP 252

QY 294 QANLVFOSIMSXGGLGLSPNLQTLFSQAXSAGARHTHTNSWGAAYGATTTDSRNDY 353

Db 253 DATLLAVRVLGPVG--SGTTEENVLAGVERAVQDADVNNISLGNLNNPDMATSTAL-DW 309

QY 354 VRKNDMTILFAAGNEXPNGGTSAPGTAKNAITVGATE---NLRPSFGSY----- 400

Db 310 AMSEGVAAVTSNGNSGRNGTVSGTSREAIISGATQPLNEVAVYFGSSAKWGVN 369

QY 401 -----ADNI----- 404

Db 370 KEDDVKALNKEVELVEAGIGEAQDFEGKDLTGKVAAYKRGSIAPVDKADNAKKAIGM 429

QY 405 -----NHVAQF 410

Db 430 VVYNNLGEIEANVPNGSVPTIKLSLEDGEKLVSAKAGETTKFTLVSKALGEQVADF 489

QY 411 SSRGPTKD-GRKPDVAPGTXILSARSSLAPDSSFANNDHSDKYAVMGTSMATPIVAGN 469

Db 490 SSRGPDVMTMKPDISAPGVNIYSTIPTHDP-----HYGYSKQGTSMASPHIAGA 543

QY 470 VAOLREHEVKNRGITPKPSL--LKAALLAGAA---DXGLGYPNGOGWGRVTLDKSLNV 523

Db 544 VAVIKQ-----AKPKMSVQGIKAIIMNTAVTLKSDGEVYPRNACAGSARI--MNA 593

QY 524 AYVNESSALSTSQAATYFTFTACKPLKISLVSDAPASTASTYL 569

Db 594 --IKADSLVSPGSYSYGTFLKENGNENKNETFTIENOSIRKSYTL 637

RESULT 7

S25835 subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain TA41)

C/Species: Bacillus sp.

C/Date: 22-Nov-1993 #sequence,revision 20-Feb-1995 #text_change 18-Jun-1999

A/Accession: S25835

R:David, S.; Feller, G.; Narinx, E.; Gerday, C.

Gene 119, 143-144, 1992

A>Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bact

A/Reference number: S25835; MUID:93012966; PMID:1398082

A/Accession: S25835

A/Molecule type: DNA

A/Residues: 1-419 <DAV>

A/Cross-references: EMBL:X63533; NID:g40198; PIDN:CAA5096.1; PID:g40199

A/Superfamily: subtilisin; subtilisin homology

C/Keywords: extracellular protein; hydrolase; serine proteinase

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-110/Domain: propeptide #status predicted <PRO>

F:111-419/Domain: microbial serine proteinase #status predicted <SBT>

F:135-373/Domain: subtilisin homology <SBT>

F:144,184,359/Active site: Asp, His, Ser #status predicted

Query Match

Best Local Similarity 9.0%; Score 279.5; DB 1; Length 419;

Matches 97; Conservative 47; Mismatches 123; Indels 61; Gaps 16;

QY 171 QXNKEVOLRGITEXIAQXXSNDVXYITAKPEYKVMNDA-----RGIVKADVAQSS 221

Db 81 QNNKRLTVEKPEL-----EIAATATNKEP--ALYNMAAASQTPWGIKATYNNNSLST 132

QY 222 YGLVGQGOIYAVVADTGIDTGRNDSMH-EARRGKITLALYALGRT--NNANDTNGHGVH 277

Db 133 SG--GAGINIAVLDTGVATNHPDLSSNVEQCKD-----FTVGNFTDNSCTDRGCHGV 185

QY 278 AGSVLGNXGTAK--GNAPQANLVFOSIM--DSXGGLGLSPNLQTLFSQAXSAGARHT 332

Db 186 AGSALANGTSGSYGVAPREADLWAKYKLGDDSGYADDIAEALRHAGDQATALTNTKVI 245

QY 333 N-SWGAAYNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGGTSAPGTAKNAITVGATE 391

Db 246 NMSLGSSESSLIT---NAVYAYDKCVLLIAAAGNSGKPSISICYPALVNAVAVALE 302
QY 392 NLRFSGSYADININVAOFFSRGPTKDG-----RIKPDVMAPTXILSARSLAPDSF 445
Db 303 NTION-GTY-----RVADFSRGRHRTAGDYIOKGDVISAFAAAYST----- 346
QY 446 MAHDSKYAYMGTSMAPIVAGNAVL 473
Db 347 W-FDGGYATISGTSMAAPHAAGLAAXI 372

RESULT 8

subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - *Bacillus halodurans*
C:Species: *Bacillus halodurans*
C/Date: 04-Dec-1992 #sequence, revision 01-Dec-2000 #text, change 15-Jun-2001
C/Accession: G83753
R:Reinhardt, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; M0ID:20512582; PMID:11058132
A:Accession: G83753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-799 <STO>
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04550.1; GSPDB:GNOC
A:Experimental source: strain C-125
C:Genetics:
A:Gene: vpr
C:Superfamily: microbial serine proteinase vpr; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:129/Domain: signal sequence #status predicted <SIG>

Query Match 8.9%; Score 276.5; DB 2; Length 799;
Best Local Similarity 23.5%; Pred. No. 1,1e-11;
Matches 199; Conservative 61; Mismatches 196; Indels 227; Gaps 25;
QY 77 LEVVPANKKLIHXPNCP-IIE-----ETKQLEXTGAKILDIYIDVYIYE 124
Db 50 LEV-----IIELEDPSTIEAHGQKQSKNELQAKQSVIEQIDLV-SSVTHE 100
QY 125 GDVAXXXXXIEHVESVEFLPYXIDPQLFTKASLVKAKALDKQXNKREVLGIEIXI 184
Db 101 YDLFESGFAL-----LPAHQIPSLIGIDGVAHVPIEYEVETDFDEV----- 144
QY 185 AOXKXSNVXYITAKPEYK-VYNDVARGIVKADVAOSSYGLGQGOIYVAVDTGLDGRN 243
Db 145 -----VIEKDAYSPFMLDSAPFICANDAMEAGY--TGEIIVAILDITGVDTHP 191
QY 244 DSSMHEAFRGKITALYAGRTNNANDT-----NGHGTAVAGSYLNGXTKKCAPO 294
Db 192 D-LVNAF-GDYKGMDFIDNNDDPQETPPDPRGIEITHTGTHVACTVANAAGLI-KGYAPD 247
QY 295 ANLYPQSLIMDSXGIGLPSMLQTL--FSQAXSAGARITITNSWGAADVGAATTTDSRVND 352
Db 248 ANLAVRYL-----GPGGAGSTAGYVAGIERAVDDGADINNLUGVTLDPDFATSL-D 302
QY 353 YKAKNDMTILFAGNEKXNGGTISAPGTAKNAITVAGATENLRPSFGSY-----ADNINH 406
Db 303 WAAABGVAAVNTSGNSGNPNMTVGSPTGSHDAISGAT--PLPKYKAKASVPTSDIDY 359
QY 407 ----- 406
Db 360 PSADIWGPPSDELELDGETEYFAFAGLGKPGDFGVDEGKIALIVGEIPEVEKAEN 419
QY 407 ----- 406
Db 420 AKAGAGVATIIYNNVAGVQPTVPGIAIPTIMLSNEDGLKMRNELBNQNTVTSIEFDKL 479
QY 407 -----VAQFSRGP-T-KDGRIPKDVMAFGTYILSARSLAPDSFANNDKSYAAWGTST 461
Db 480 VGETVADFSRGRHRTAGDYIOKGDVISAFAAAYST-----GYGSRQSTMS 533

QY 462 APPIVAGNAOLREHFVKNRGTTPKPSLLKALIAAGAA-----DXGLGYPNGOGWRVTL 517
Db 534 ASHVAGAAALLLEAH-PNMGV-----DHVKAALMTAEVLVDENGHRPHNTOGAG----- 584
QY 518 DKLNAVAVNNESSLSTOKATY-XFATAGKPLK 551
Db 585 --SIRIVAISETLVTPGSHSFGTFTKGRKQVE 617

RESULT 9

subtilisin (EC 3.4.21.62) 1 precursor - *Bacillus* sp. (strain TA39)
S23407
C:Species: *Bacillus* sp.
C/Date: 04-Dec-1992 #sequence, revision 04-Dec-1992 #text, change 18-Jun-1999
C/Accession: S23407
R:Marinx, E.; Davall, S.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1131, 111-113, 1992
A:Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
A:Reference number: S23407; M0ID:9256481; PMID:1581352
A:Accession: S23407
A:Molecule type: DNA
A:Residues: 1-420 <NAB>
A:Cross-references: EMBL:X62369; NID:g40200; PIDN:CAA44227.1; PID:g40201
C:Genetics:
A:Gene: subtil
C:Superfamily: subtilisin; subtilisin homology
C:Keywords: extracellular protease; hydrolase; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-111/Domain: propeptide #status predicted <PRO>
F:112-420/Product: microbial serine proteinase #status predicted <MAT>
F:136-374/Domain: subtilisin homology <SBR>
F:145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 8.9%; Score 274.5; DB 1; Length 420;
Best Local Similarity 28.3%; Pred. No. 6.3e-12;
Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;
QY 167 LPTKQXN-KEYQLRGIEIXIAQXXXXNXYITAKPE-YKVM--NDVARGIV---KA 215
Db 73 MMEKQFMALKRKKMLTYEKVPEL---EIAATDTRFALYNNMAASQTPWGIKATYNNNS 128
QY 216 DVAOSSYGLYGGOIVAVADTGLDGRNDSMHEAFRGKITAL--YALGRT--NNANDT 270
Db 129 SIQRTS---GGGGINIVAVLDITGVNTN-----HPDLRNVECKDFVTGTYTNNSCDR 179
QY 271 NGHGTAVAGSYLNGXTNKK---GMAPOANLVFOSTW--DSXGGLGGLPSNLQTLFSAQXS 325
Db 180 OGHGTAVAGSALADSGTNGVGYGAAPDADLMAYKVLGDDGSGVADTAAAIIRHAGQDATA 239
QY 326 AGARITHTN-SMGAAVNGAYTTDSRNDVYRKNDMTILFAAGNEKXNGGTISAPGTAKNA 384
Db 240 LNTKRYINSLGSSSESLITNAVN---YSNKGVLITIAAGNSGPTQSGISGIPGALVNA 296
QY 385 ITVGATENTLRPSFGSYADININVAOFFSRGPT-KDG-----RIKPDVMAPTXILSARSS 438
Db 297 VAAVALEN-KVENGY-----RVADFSRGRHRTAGDYIOKGDVISAFAAAYST--- 347
QY 439 LAPDSFANNDKSKAYVWGTSMATPIVAGNAOLREHFVKNRGTTPKPSL-----L 490
Db 348 -----W--FDGGYATISGTSMAAPHAAGLAAKIWAQYPSASNVVAGELQYRAVENDI 398
QY 491 KAALTAGAAD---XGLGY 505
Db 399 LSGYYAGYGDDEFASGQGF 416

RESULT 10

S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - *Xanthomonas campestris* pv.
N:Alternate names: subtilisin-related proteinase
C:Species: *Xanthomonas campestris* pv. *campestris*
C/Date: 21-Nov-1993 #sequence, revision 07-Feb-1997 #text, change 03-Dec-1999

C/Accession: S11890
C/Date: Y.N.: Tanq, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A/Title: A multipurpose broad host range cloning vector and its use to characterise an
A/Reference number: S11890; MUID:90251253; PMID:2187155
A/Accession: S11890
A/Molecule type: DNA
A/Residues: 1-580 <LNU>
A/Cross-references: EMBL:X51635; NID:948533; PIDN:CAA35962.1; PID:948534
A/Experimental source: Xanthomonas campestris pv. campestris
A/Note: The sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-Alt
C/Superfamily: subtilisin; subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine proteinase
F:1-32/Domain: signal sequence #status predicted <SIG>
F:168-423/Domain: subtilisin homology <SBT>

Query Match 8.8%; Score 273.5; DB 2; Length 560;
Best Local Similarity 22.7%; Pred. No. 1.2e-11;
Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;

QY 1 MRKKVFLSVSAALITVALNMPAGXARXPDLPFGTCTTDDXGFSKQXTGAAF 60
D 6 LKRTGSLTIGASALTSLAMPAG---EYIDGLTAQTHQKFTVYADGSL 60
QY 61 LVESENKXKGLKXKLETPANKLHXDPNGPILBETKXLEXTGAKILDYIPDYAI 120
D 61 ASPS---ALTTSLRTAARAPA-----KAGKALGINSVRLALGPELV-- 100
QY 121 VERGDVXKXXXIEHVESVEPYLPXYXIDPQLFTKASLXVKAALDTKXKNEVOLRG 180
D 101 -----RADRALDRAELETMLRDLADPNVQSEV----- 129
QY 181 IEIXAQXXSNDVXYITAKP-EYKVMNDVARGIYKA--DVAKSSYGLGOGQIYVAVADG 237
D 130 -----DQILHATLPDTRLSEQWAFGTTNAGINIRPADKAGSSTVAIVDTG 179
QY 238 L---DTGRNDSMHEAFRGKTALYALGRTNAND-----TNGHGT 275
D 180 ITSHADINANILAGYDFSDATPTARDGSRDSDNADEGDWYAADECAGIPAASSSMHGT 239
QY 276 HVASV---LGNGXKN-KGMAQANLVFQSIDMSXG-----LGGLPBNL 316
D 240 HVAQTVAAVTNTGTGAGTAIGAKVVPVRYLGKCGSLSDIADAIWASGGTVSGIIPNA 299
QY 317 Q--TLFSAQXAGARHTNSGAAVNGAYTTDSRNVDYVRKNDMTILFAENEXPNCGT 374
D 300 NPAYVIMMSLGGGSGCSTTQGN-ALNGAVSNGT-----TVYVAGNDASNVSG 346
QY 375 ISAGTAKNAITVGATEN--LRPSFGSYADNINHVAQESSRGPTRKGRIRKPDVNAFGTXI 432
D 347 -SLRANCANVAIVAAITTSAGAKASYSNFGTGI-----DVSAFGSSI 386
QY 433 LSARSS--LAPDSSFMANHDSKAYVMGTSMAPTPIVAGNVOLREHPFKNGITPK--PS 488
D 387 LSTINSGTTPGS-----ASTASYNGTSMASPHVAGVALVOS--VAPPALPAAVET 437
QY 489 LK--AALIIAGADXXGLGY-----PNGNOGWRVTLDKSLVAVAVNESSX 531
D 438 LKMTARALPACSGGCGAGIVNDAVTAALINGSGGCGGGMTLTNGTPTVIGL----- 493
QY 532 LSTQKATYATATAGAPKLSIWSDAFSTASVTL--VNDLDLY--TAPNGTXY 585
D 494 AATCAELNLTITVPAG-----SGTLVTTSGGSGDADLYVRAGSAPTDSAY 539
QY 586 VGNDFFXPKXXNMGRNNVENVFNPQSGTYIEVOAYN 625
D 540 TCRPYRS-----GNAETCTITAP-SGTYIVRLKAYS 569

RESULT 11
JC4908
alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
N/Alternate names: subtilase

C/Species: Alteromonas sp.
C/Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
C/Accession: JC4908
R/Tsujiho, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y
BioSci. Biotechnol. Biochem. 60, 1284-1288, 1996
A/Title: Cloning and sequence analysis of a protease-encoding gene from the marine ba
A/Reference number: JC4908; MUID:97141200; PMID:8987544
A/Accession: JC4908
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-715 <TSU>
A/Cross-references: DDBJ:D38600; NID:91536787; PIDN:BAI18912.1; PID:d1019647; PID:g21
A/Experimental source: strain O-7
C/Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-se
C/Genetics:
A/Gene: aprI
C/Superfamily: subtilisin homology
C/Keywords: hydrolase
F:1-40/Domain: signal sequence #status predicted <SIG>
F:41-150/Domain: amino-terminal propeptide #status predicted <AMP>
F:151-496/Product: alkaline serine protease I #status predicted <MNT>
F:182-452/Domain: subtilisin homology <SBT>
F:457-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 8.7%; Score 268.5; DB 2; Length 715;
Best Local Similarity 25.8%; Pred. No. 3.5e-11;
Matches 142; Conservative 51; Mismatches 201; Indels 157; Gaps 28;

QY 168 DTGXKNEVOLRGIEIXAQXXSNDVXYIT-----AKPEYKVMND-----VARGIV 213
D 117 DKLSAKEQAQAFNEVVA---SGNVEYIEIDQMLKFPAPNDPRYDQWHYBAAGI- 171
QY 214 KADVAQSSYGLGOGQIYVAVADG---LDYGRNDSMHEAFRGKTALYALGRTNAND 269
D 172 NAPAAMDK--ATGGGVVAVVLDYGRHLDLDAVILPEYDITSTFYVANDGARDNDARD 229
QY 270 ---TNG-----HGTNVAG---SVLNGXKTKGAPQANLVFQS 301
D 230 PGDAVTRGEGCTDSSGQVPVRADQDSSMHGTHVAGTAAVTNNGEVAGVAYDAKVVR 289
QY 302 IMDSXGLIGLPSNLQTLFSAQXAGARHTNSGAAV-----NCAVTTDSRNVDYV 354
D 290 VL---GKCGGLTSDIADGIIWASGSDRVPANAPVAVINMSLGGGACSAATTQMAINO 346
QY 355 RKNMTLIFAGNEXPNCGTISAPGTAKNATVGAT--ENLRPSFGSYADNINHVAQESS 412
D 347 RNNGTVIYIAAGNDNDSANIN-PGNCNGVYVAVASVGRDGRATYSNGANI----- 397
QY 413 RGPTKGRIRKPDVNAFGTXIILSARSSLAPDSSFVANHDS-----KYAVMGTSNATP 464
D 398 -----DVAAPG---GAQSFADDEPGLTSTHNSGSGAPSDNSYHSGQTSMAAP 442
QY 465 IVAGNVQLR-----EHFVKN--RGITPKPSILKRALI--AGAADXGIG-----Y 505
D 443 HVAGVAAALIKQAKSPATPEDEVETIKNTTRSFACSCNCGTGVDAAAVANALGVVTP 502
QY 506 PNGNOGWRVTLDKSLVAVAVNESSXLSOKATYXFTATAGKPLKLSIWSDAFSTATA 565
D 503 PTGN-----TLED--GVAKTGLSGAAGSND--FTTVDVRAK-----TNV 538
QY 566 SVTL---VNDLDVITAPNGTYVGNDFXPKXXNMGR---NNVENVFNPQSGTYT 618
D 539 TETWSGSGTADADLYVK-----LGSG---PTSSSYDCRPREGCNALVCSFADPAQACTYH 588
QY 619 IEVOATVNPVG 629
D 589 VMINGYKAYS 599

RESULT 12
F69730
cell wall-associated protein precursor wpra - Bacillus subtilis

C:Species: *Bacillus subtilis*
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: F69730
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta
 C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emerson, P.T.; Estlin, K.D.; Errington, J.; Farrell, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
 A.; Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A.; Laubert, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Roche, B.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Stanton,
 A.; Authors: Schleicher, S.; Schuster, R.; Scofield, P.; Sekiguchi, J.; Sekowsky, A.; Serot
 A.; Authors: Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, E.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: F69730
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-894 <KUN>
 A:Cross-references: GB:299109; GB:AL009126; NID:92633260; PIDN:CAB12917.1; PID:el183079;
 A:Experimental source: strain 168
 A:Genetics:
 A:Gene: wprA
 A:Superfamily: subtilisin homology
 F:453-664/Domain: subtilisin homology <SBT>
 Query Match 8.4%; Score 262; DB 2; Length 894;
 Best Local Similarity 23.9%; Pred. No. 1.4e-10;
 Matches 134; Conservative 92; Mismatches 21; Indels 124; Gaps 28;

Oy 49 FSKOXOTGAAP--LVESENVKLLKGLKKLETPA-----NNKLLHXQNP---G 93
 Db 245 FDDVSENGASSYKTEKDKAINRLDKALQSVFLKEIKKQADRLNKKQLOQKTAG 304
 Oy 94 PILEE-----TKOXLETKAKLIDYIPYAYIV---ETEG-----DYXSSXXX 133
 Db 305 ALTEENNAKSEVOIT--KTFKDKKSLSSVHNHMKGSFASAKSKKIDSNVAKAKL 362
 Oy 134 IEHVESVEPELYXYIDPOLFTKQASKLVAKALDTKQKNKEVQRLGIEIXIAQ--XXXND 192
 Db 363 FDMUYSEF--LPRDEKONGAYTASAKRVKSAALTLKMSNVEF---AEPQEKSLAND 416
 Oy 193 VXYITAKPEYKVMNDYARGIVK-ADVA--QSSYGLYQCOG---IYAVADTGLDTRNDSS 246
 Db 417 IOY---PYQMPKLNNGENGCVKADVKEYPANTLLSKRKLNDTLAIVDVGDSFLAD-- 471
 Oy 247 MHEAFRGKITALYA---LGRTNANDTNGHGTIVAGSYLG---NGXTNKGMAPQANLVPQ 300
 Db 472 ----LKGVRDTLGHNFVGRNNNAMDQGHGTIVAGIIAAGSDNGYSMTGLAKAKIIPV 527
 Oy 301 SIMDSXGGLGGLPSNLQTLFQSAASAGARLHTNSMCAAVNCAVTTDSRNVDYRRKNDMT 360
 Db 528 KVIDSAG--SGDTQIALGIVKADKAKINISLG---GGYKRVLEFALKADKAVL 561
 Oy 361 ILPAANEKPNNGTISAPGTAKKAIYIGATEINLRPSFGSYADININHVAFSSRGPTDGR 420
 Db 582 IAAASGNDEN--ALSTPSSKIVYMSVGA-- --NMDDMTADFNSYKGL-- 625
 Oy 421 IKPDVAPGTITLSARSSSLAPDSSFFMANHDSKYAVMGCTSMATPIYAGNVQLREHFVN 480
 Db 626 ---DISAPSDT---PSLIVNGN-----VTYMSGTSMATPYAAAAAGTLFQAQPKL 670
 Oy 481 RGITPPPSLAKA--LIAGADG-----LGP-----NGNCGMRVLTLSLN 522
 Db 671 KRTEVEDMKKTADDISFESVDGEEELDYDGPDIETPKTPGVWMHSGYGLNVMKAVS 730
 Oy 523 VA-----YVNESSXLSQKAT 539
 Db 731 AADQLKVNKLSTQTAVRGS 751

RESULT 13
 T29090
 Surface layer-associated STABLE proteinase - *Staphylothermus marinus*
 N:Alternate names: hyperthermostable proteinase
 C:Species: *Staphylothermus marinus*
 C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C:Accession: T29090
 R:Mayr, J.; Lupas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.
 Curr. Biol. 6, 739-749, 1996
 A:Title: A hyperthermostable protease of the subtilisin family bound to the surface
 A:Reference number: Z20559; MUID:96385442; PMID:8793300
 A:Accession: T29090
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1345 <MAY>
 A:Cross-references: EMBL:U07968; NID:91374755; PID:91374756; PIDN:AAB02323.1
 A:Experimental source: strain F1
 A:Function:
 A:Note: Stochastic S-layer component
 Query Match 8.4%; Score 260.5; DB 2; Length 1345;
 Best Local Similarity 29.1%; Pred. No. 3e-10;
 Matches 98; Conservative 41; Mismatches 105; Indels 93; Gaps 14;

Oy 251 FRGKITALYALGRTNANDTNGHGTIVA-----GSVL-----GNGXTNK--GNAPQANLV 298
 Db 445 YQRYLAL-----YDFPHGQTSVATVYASRGVLYLDYGCKLYRIMGVAPGAKI- 495
 Oy 299 FQSIMDSXGGLGGLPSNLQTLFQSAASAG-----EAMLAGENITREEDGYVLSLDPPGPHRIDISNSM 335
 Db 496 -----ACGDAMLLGNILVL--EAMLAGENITREEDGYVLSLDPPGPHRIDISNSM 546
 Oy 336 GAANVATYTDNRND-----DYRKNNDMTILPAAGKEXPNGTISAPGTAK 382
 Db 547 GSTYINFWLODPGIDTRSSFMDEIIAIRNLTIGDVTITVPAAGKSGCSSNCAPEQGL 606
 Oy 383 NAITVGATE--NLRFSG--STADNINHVAFSSRGPTKDGRIKPDVMAFG-----TX 431
 Db 607 LVTTACASTLMDYTRIGYPCGVAD--EVTPESSRGPTGOQYPPDVTIVGAFENASTR 663
 Oy 432 IISARSLAPDSSFFMANHDSKYAVMGCTSMATPIYAGNVQLREHFVNKRGITPKSLK 491
 Db 664 TIDRGYGAQDPVF-----GQTSERFPTSGTLALVFAQYEVYVTPDPVPAK 712
 Oy 492 AALIAAGADGXLGYPNGNGMGWRTLDKSLNVAAYNE 528
 Db 713 IILKSSAND--IWPAPFQSGGRVDAKADVPYFISE 747

RESULT 14
 A33973
 High-alkaline serine proteinase (EC 3.4.21.-) yab precursor - *Bacillus* sp. (strain ya
 N:Alternate names: alkaline elastase; subtilisin yab
 C:Species: *Bacillus* sp.
 C>Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 22-Jun-1999
 C:Accession: A33973; B33973
 R:Kaneko, R.; Koyama, N.; Tsal, Y.C.; Juang, R.Y.; Yoda, K.; Yamasaki, M.
 J. Bacteriol. 171, 5212-5236, 1989
 A:Title: Molecular cloning of the structural gene for alkaline elastase yab, a new su
 A:Reference number: A33973; MUID:89359181; PMID:2670913
 A:Accession: A33973
 A:Molecule type: DNA
 A:Residues: 1-378 <KAN>
 A:Cross-references: GB:M8537; NID:9341960; PIDN:AAA87324.1; PID:9758666
 A:Note: parts of this sequence, including the amino end of the mature protein, were d
 A:Accession: B33973
 A:Molecule type: Protein
 A:Residues: 111-164;326-355 <KA2>
 A:Genetics:
 A:Gene: ale

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:28:13 ; Search time 9.99218 Seconds

(without alignments)
2652.411 Million cell updates/sec

Title: US-09-920-954-1

Perfect score: 3101
Sequence: 1 MKKKVSLVLSAAAILSTV.....EVOAVNPVGQKFSIAIVN 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518.5	16.7	1743	1	TAGC_DICDI
2	477.5	15.4	1905	1	TAGC_DICDI
3	299.5	9.7	1398	1	PLS_PYRPU
4	286	9.2	806	1	SUBV_BACSU
5	274.5	8.9	420	1	SUBT_BACSP
6	273.5	8.8	580	1	PROA_XANCP
7	262	8.4	894	1	MPRA_BACSU
8	257.5	8.3	378	1	ELVA_BACNA
9	255.5	8.2	381	1	SUBN_BACNA
10	255	8.2	513	1	NOI_THRNO
11	253	8.2	382	1	SUBT_BACNA
12	251.5	8.1	381	1	SUBT_BACNA
13	250.5	8.1	380	1	ELVA_BACNA
14	250.5	8.1	380	1	ELVA_BACNA
15	249.5	8.0	379	1	SUBT_BACLI
16	249.5	8.0	381	1	SUBT_BACSU
17	247.5	8.0	402	1	SUBT_BACST
18	246.5	7.9	401	1	ALP_CEPAC
19	239.5	7.7	534	1	THES_BACSP
20	239.5	7.7	269	1	PROA_VIBAL
21	239	7.7	269	1	SUBS_BACLE
22	238	7.7	269	1	PRTM_BACSP
23	234.5	7.6	1181	1	SCA2_STRPY
24	232	7.6	1181	1	SCA2_STRPY
25	231.5	7.5	1167	1	SCAL_STRPY
26	223.5	7.2	404	1	SMP1_MAGPO
27	223.5	7.2	409	1	ALP_TRIHA
28	223	7.2	321	1	ISP_BACCS
29	218.5	7.0	275	1	SUBT_BACPU
30	218.5	7.0	533	1	PEPC_ASPNG
31	218	7.0	361	1	ELVA_BACHD
32	216.5	7.0	645	1	SUBE_BACSU
33	214.5	6.9	326	1	ISP_PAEPO

34	214	6.9	293	1	PRTT_TRIAL	P20015 tritirachlu
35	212.5	6.9	274	1	SUBD_BACLI	P00781 bacillus 11
36	212	6.8	603	1	BPRV_BACNO	P42779 bacteroides
37	211.5	6.8	1052	1	MSIP_CRIGR	O9228 cricetus
38	211.5	6.8	1052	1	MSIP_MOUSE	O9228 mus musculus
39	211.5	6.8	1052	1	MSIP_RAT	O9228 rat
40	211.5	6.8	1433	1	SUBP_BACSU	P16397 bacillus su
41	211	6.8	530	1	HLV_HALL17	P29143 halophilic
42	210.5	6.8	1052	1	MSIP_HUMAN	O14703 homo sapien
43	210	6.8	422	1	TKSU_PYRNO	P58502 pyrococcus
44	210	6.8	595	1	DPRK_BACNO	P42780 bacteroides
45	209	6.7	408	1	SEPR_THESR	P80146 thermus sp.

ALIGNMENTS

RESULT 1	ID	TAGC_DICDI	STANDARD:	PRT: 1743 AA.	
AC	Q23868	dictyostell			
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Presalk-specific protein tag precursor (EC 3.4.21.-).				
GN	TAGC				
OS	Dictyostellum discoideum (Slime mold).				
OC	Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.				
OX	NCBI_TaxID=4689;				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=AX4;				
RX	MEDLINE=97140317; PubMed=8986798;				
RT	Shaulsky G., Escalante R., Loomis W.F.;				
RT	"Developmental signal transduction pathways uncovered by genetic				
RT	suppressors."				
RT	Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).				
CC	- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGC MAY MEDIATE				
CC	INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS (BY				
CC	SIMILARITY).				
CC	- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO PEPTIDASE FAMILY				
CC	S8.				
CC	- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING				
CC	TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.				
CC	- SIMILARITY: STRONG, TO TAGC.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: U60086; AAB0331.1; -				
DR	DictyDb: DD02060; tagc.				
DR	InterPro: IPR003593; AAA_ATPase.				
DR	InterPro: IPR003439; ABC_TRANSPORT.				
DR	InterPro: IPR001140; ABCtransportTM.				
DR	InterPro: IPR000209; Peptidase_S8.				
DR	Pfam: PF00005; ABC_tran; 1.				
DR	Pfam: PF00664; ABC_membrane; 1.				
DR	PRINTS: PR00723; SUBTILISIN.				
DR	SMART: SMO0336; SUBTILASE ASP; FALSE_NEG.				
DR	PROSITE: PS00137; SUBTILASE_HIS; 1.				
DR	PROSITE: PS00139; SUBTILASE_SER; FALSE_NEG.				
DR	PROSITE: PS00211; ABC_TRANSPORTER; 1.				
KW	Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;				
KW	Signal.				
FT	STGNL	1	27		POTENTIAL.
FT	CHAIN	28	1743		PRESTALK-SPECIFIC PROTEIN TAGC.
FT	DOMAIN	316	642		PROTEASE.


```

FT DOMAIN ? 1743 ABC TRANSPORTER.
FT TRANSMEM 962 982 POTENTIAL.
FT TRANSMEM 1027 1047 POTENTIAL.
FT TRANSMEM 1072 1092 POTENTIAL.
FT TRANSMEM 1157 1177 POTENTIAL.
FT TRANSMEM 1260 1280 POTENTIAL.
FT TRANSMEM 1288 1308 POTENTIAL.
FT ACT_SITE 325 325 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 637 637 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1485 1492 ATP (POTENTIAL).
FT DOMAIN 42 46 POLY-ASN.
FT DOMAIN 94 103 POLY-ASN.
FT DOMAIN 643 646 POLY-ALA.
FT DOMAIN 733 741 POLY-ASN.
FT DOMAIN 786 792 POLY-SER.
FT DOMAIN 1337 1340 POLY-GLU.
FT DOMAIN 1346 1352 POLY-GLY.
FT DOMAIN 1353 1357 POLY-ASN.
FT DOMAIN 1358 1364 POLY-ASP.
FT DOMAIN 1381 1386 POLY-ASN.
FT DOMAIN 1707 1729 POLY-ASN.
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 689 689 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 741 741 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 832 832 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1251 1251 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1385 1385 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1386 1386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1454 1454 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1704 1704 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1743 AA; 194145 MW; 12DB363E2F729839 CRC64;

```

Query Match 16.7% Score 518.5; DB 1; Length 1743;
 Best Local Similarity 24.2%; Pred. No. 4.1e-30;
 Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

```

QY 87 HIXQFNGPILETROKLT-----EXTGAKIDYIPDVAIVY--EYEGDVXSK 130
DQ 115 YIVQFKDINDETRQLEKFLIGTDIVDEQYQSHIVHYIPHDSFLVLMQEQSVLLSS 174
QY 131 XXXIEHVSVEPY-----LPXYXIDPOLFTKGASXLVK--AXALDT--KQXN 173
DQ 175 KEWMSWIGEFEPSPKRIHLNNEKSGILPVIYI---LSDSTNSLIQRMENPLNSILKSYN 230
QY 174 KEVOLRGI-----EXIAQXXXSDVXYITAKKPEYK 203
DQ 231 SKVLTILINQKKLSIYVNCNDESPSCSLINSEKLYQWISQSESN---FIESEKFFQ 287
QY 204 VMDVARGIV-----KADVAQSSYGLYGQGLVAAVDTGLDGR--NDS----- 245
DQ 288 TANRISPRVVGCTKDTLVNNDVRDIP-----LRGKGQLTSLADTGLDSHCFQFSQXPI 342
QY 246 ---SMHEAFKRIITALYALGRTNNANDTNGCHTHVAGSVLC-----NGXTKNGMAPOA 295
DQ 343 PLNSVNLNHR-KVYVYITTSDDSDKVDYGHGHTHICGSAAGTPEDSSVNISFSGLATDA 401
QY 296 NLVQSIWDSXGGLGL--PENLQTLFQSAQXAGARIFHTNMGW-----AVNGAYITDSRN 349
DQ 402 KIAF--FDLASGSSSLTPPSDLKQLYOPLVADAGARVCDMSGVSVEGYTSSDTSAS 458
QY 350 VDDYVRKN-DMTILFAAGNEXPNGGTIS--APGTAKNAITGATENLR-----PSRGS 399
DQ 459 IDDFLETHPDIILRAAQN---NEQYLSLLQSTAKNAVITYGAGHTIHENLITDGPNTIN 515
QY 400 YADNT----- 404

```

```

DQ 516 YQSSVDINQELICDPDSRYCNYTTAQCCLSNATTTGLASCPILLKRSVIDAANTOPLY 575
QY 405 --NHVAQFSRGPTRKDRKIPDVNAPGTXLISARSSLA-----PDSSFMANHDSKYAYM 456
DQ 576 NENNICFSSSGKPHDQRMKRALVAPGEYITSAANGCANMTDQCGDSL--PNTALLA-I 633
QY 457 GGTSMAPIYAGNVAQUREH-----FVKRGITPRPSLKALALIGA----- 498
DQ 634 SGTSMASFAAAMATILROYLVQGYPTGSLVESNKLQPPGSLKALMINNAQLLNGTFQ 693
QY 499 --AAXGIGYNGN-----QGWGRVYTDKSLNVAAYNESS----- 530
DQ 694 LITSSSITPSNOVFENFACASLVQMGALIMSMWLVHVNNNNNNNKTSDDGTRKPDGI 753
QY 531 -----XLTSTQKATYXT-----ATAGKPLK--ISLWMSDAPASTT 564
DQ 754 GGLDLRLVKPMQKKEESLSTQNTSYCFYTPSSSSNSGNNIPRVATLVMTDPPSYAG 813
QY 565 ASVTLVNDLIVT-----TAPNGTYXGNDPFXAPXXXNMGRNVENVFTINXP 612
DQ 814 AKFMLVNNLDTMTIYRDNGSTIFYSNQGGSSFLG-----LAPTDPLNNVEGIVHNP 867
QY 613 QSGTYTIEVQAYNPVGPQXFS 634
DQ 868 EPMTRYEMVACTNVPMPQPNFS 889

```

RESULT 2

ID TAGB_DICDI STANDARD; PRT; 1905 AA.

AC P54683;

DT 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-1996 (Rel. 34, Last sequence update)

DE Prestalk-specific protein tagB precursor (EC 3.4.21.-).

GN TAGB.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

NCBI_Taxid=44689;

RN [1]

RC STRAIN=AX4;

RA MEDLINE=95262903; Pubmed=7744252;

RT Shaubsky G., Kuspa A., Loomis W.F.;

RT "A multidrug resistance transporter/serine protease gene is required for prestalk specialization in Dictyostelium.";

RL Genes Dev. 9:1111-1122(1995).

CC -1- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGB MAY MEDIATE INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO PEPTIDASE FAMILY S8.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.

CC -1- SIMILARITY: STRONG, TO TAGC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U20432; AAA62212.1; -.

DR MEROPS; S08.0PW; -.

DR Dictydb; DD02059; tagB.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003439; ABC_transporter.

DR InterPro; IPR001140; ABCtransprtTM.

DR InterPro; IPR000209; peptidase_S8.

DR Pfam; PF00005; ABC_tran.1.

DR Pfam; PF00664; ABC_membrane.1.

DR PRINTS; PR00723; SUBTILISIN.


```

DR ProDom: PD000006; ABC-transport: 1.
DR SMART: SM00382; AAA: 1.
DR PROSITE: PS00136; SUBTILASE-ASP; FALSE-NEG.
DR PROSITE: PS00137; SUBTILASE-HIS; 1.
DR PROSITE: PS00138; SUBTILASE-SER; 1.
DR PROSITE: PS00211; ABC-TRANSPORTER; 1.
KW Hydrolyase; Serine protease; ATP-binding; Transport; Transmembrane;
KW Signal.
FT SIGNAL. 1 31
FT CHAIN 32 1905
FT DOMAIN 378 700
FT TRANSMEM 1011 1031
FT TRANSMEM 1076 1096
FT TRANSMEM 1121 1141
FT TRANSMEM 1210 1230
FT TRANSMEM 1309 1329
FT TRANSMEM 1332 1352
FT AC1_SITE 387 387
FT AC1_SITE 432 432
FT AC1_SITE 432 432
FT AC1_SITE 695 695
FT NP_BIND 1553 1560
FT DOMAIN 63 67
FT DOMAIN 95 104
FT DOMAIN 107 134
FT DOMAIN 311 321
FT DOMAIN 833 837
FT DOMAIN 838 844
FT DOMAIN 871 876
FT DOMAIN 1012 1015
FT DOMAIN 1386 1389
FT DOMAIN 1445 1450
FT DOMAIN 1765 1778
FT DOMAIN 1782 1785
FT DOMAIN 1807 1812
FT DOMAIN 1813 1860
FT CARBOHYD 1872 1878
FT CARBOHYD 594 594
FT CARBOHYD 621 621
FT CARBOHYD 672 672
FT CARBOHYD 747 747
FT CARBOHYD 823 823
FT CARBOHYD 1172 1172
FT CARBOHYD 1522 1522
FT CARBOHYD 1658 1658
SQ SEQUENCE 1905 AA; 212518 MW; B8E23F8AB9A13C CRC64;

Query Match 15.4%; Score 477.5; DB: 1; Length 1905;
Best Local Similarity 24.5%; Pred. No. 4; 8e-37;
Matches 193; Conservative 98; Mismatches 237; Indels 259; Gaps 31;

OY 87 HIXQFNPILIEETKQXLE---XTGAKLIDYIPDYAVIYER-----BDYXS 129
DB 172 YIVQFKRINDETREQLKEFLIGTITLKEQPKSHIVHIFHDSFLVMTKEQSVLLS 231
OY 130 XXXXIEHVESPEY-----LPXXYI---DPQLFTKGASXL-----V 162
DB 232 SKEMISWIGEHPKMKILNLYHEKSGIPVYIILSGSTNLIQRMWNTLSILTSYNSKV 291
OY 163 KAKALDTRQ-----XNKEVQLRGIEIXIAQXXS--NDVYIITAKPEYVM 205
DB 292 KLTLINRKLKLSIYCNDESPSSSSSSCSLIGSEKIVYKWISEGESNYIERSEKLTQA 351
OY 206 NDVARGIV-----KADVAOSSYGLGOGQIYAANDTGIDGR---NDS----- 245
DB 352 NRLSPFVIFETKDKLVNDRIDIP-----LRKGQGIISINDTGLDSDSHCFSDSKYPIPF 406
OY 246 SMHEAFRGKITALVLTGRTNANNDTNGHGVAGSVLGKGN-----KGMAPQANL 297
DB 407 NOVVENHRRKVTY---YITTHDEVDVYNGHGHVCCSAAGPEPDSNAISSFGSLATDKTI 463
OY 298 VFOSIMDSXGGLGCLPSNLQTLFSAQXASGARIFHTNSWGA-----AVNGAYTTDSRNVVDY 353

```

```

DB 464 AFYD-LSSGSSEPTPEDEYSQMYKELYDAGARVHGDSWGSVSLGGYGGSDAGGIDAF 522
OY 354 VRK-NDMTILFAAGNEXPRGCTISAPGAKNAIIVGATEULRPSFGS-----YADNI- 404
DB 523 LYEPERSLIRAAAGNN-ELFASLAAQATAKNAIIVGAEQFAHNVYSDALEYDFPSDMAN 581
OY 405 -----NHVAQF 410
DB 582 FQRCLEFDKRYCYNTTAKCCSEVSNVKGLOLCCPASIKONASDFTTORPOTENNNKSF 641
OY 411 SSRGPTKDGKIKEDVWAPGTXILSARRS-----LAPDSFMANHDSKAYVGGTSNATP 464
DB 642 SSKGPTHGDKEDVAPGEXYTSARSNGENSTDOCGDSL--PNANGIMLSIGTSMATP 699
OY 465 IVAGNVAQLREHF-----YKNRGITPPKPSLLKALI----- 495
DB 700 LATATATILRQYLVGDYFPPTGESVENKLLPTGSLKALMINNAOLLNGYFWSASSTNP 759
OY 496 AGADKXGLGYPNGQGMGRVTLDKSLNVAVYNESS-----XLISTOKAT----- 539
DB 760 SMAIFEQINGANLIQGMGLRWN--NMLYKSSNPTPPSRMIGIGLGKKNOKATEWKED 816
OY 540 -----YXFT-----ATAGKP-LKISLWSDAPASTASTVLYNDLDL--- 575
DB 817 SLSSGLNKSYCFIYKPSRSSSSSGSGGGTPTVATVLTWDPSPYSAKRNLVNNDLILL 876
OY 576 -----YITAPN--GTYXYGNDFFXXPXAMMDGNVENVFINXQSGTYITLVOAVNVP 627
DB 877 NSDDSLITIGNSGSLQPAKVAP-----DTLNVEGIIINPTKANNYKFTTAGTNVP 931
OY 628 VGPQXES 634
DB 932 IGPOKFS 938

RESULT 3
PUS_PYRFU STANDARD; PRT; 1398 AA.
ID PUS_PYRFU
AC P72186;
DT 16-OCT-2001 (rel. 40; Created)
DT 15-JUN-2002 (rel. 41; Last sequence update)
DT 15-JUN-2002 (rel. 41; Last annotation update)
DE Pyrolysin precursor (EC 3.4.21.-).
GN PUS OR PF0287.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 150-184, AND CHARACTERIZATION.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=9635370; PubMed=8702780;
RA Voorthorst W.G.B., Eggen R.I.L., Geerling A.C.M., Platteeuw C.,
RA Slezan R.J., de Vos W.M.;
RT "Isolation and characterization of the hyperthermostable serine
RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
RT Pyrococcus furiosus.";
RL J. Biol. Chem. 271:20426-20431(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RN Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
RX CHARACTERIZATION, AND 3D-STRUCTURE MODELING.
RX MEDLINE=21079021; PubMed=11210516;
RA de Vos W.M., Voorthorst W.G.B., Dijkgraaf M., Kluskens L.D.,
RA Van der Oost J., Slezan R.J.;
RT "Purification, characterization, and molecular modeling of pyrolysin
RT and other extracellular thermostable serine proteases from
RT hyperthermophilic microorganisms.";

```

RL Meth. Enzymol. 330:383-393(2001).
 CC -1- FUNCTION: Has endopeptidase activity toward caseins, casein
 CC fragments including alpha-S1-casein and synthetic peptides.
 CC -1- SUBCELLULAR LOCATION: Cell-envelope associated.
 CC -1- PTM: LMN pyrolysins seems to be produced by autoprotoeolytic
 CC activation of HMW pyrolysin.
 CC -1- PTM: Glycosylated.
 CC -1- MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, U55835; AAB09761.1; -.
 DR EMBL, AE010153; AAL80411.1; -.
 DR HSSP, Q45670; IDB1.
 DR MEROPS: S08.100; -.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 4.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 DR Hydrolyse: Serine protease; Zymogen; Signal:
 KM Complete proteome.
 KW Complete proteome.
 FT SIGNAL 1 ?
 FT PROPEP 1 149
 FT CHAIN 150 1398
 FT ACT_SITE 179 179
 FT ACT_SITE 365 365
 FT ACT_SITE 441 441
 FT ACT_SITE 441 441
 FT CARBOHYD 152 152
 FT CARBOHYD 222 222
 FT CARBOHYD 228 228
 FT CARBOHYD 240 240
 FT CARBOHYD 257 257
 FT CARBOHYD 262 262
 FT CARBOHYD 298 298
 FT CARBOHYD 327 327
 FT CARBOHYD 406 406
 FT CARBOHYD 406 406
 FT CARBOHYD 651 651
 FT CARBOHYD 663 663
 FT CARBOHYD 739 739
 FT CARBOHYD 792 792
 FT CARBOHYD 893 893
 FT CARBOHYD 908 908
 FT CARBOHYD 917 917
 FT CARBOHYD 929 929
 FT CARBOHYD 1048 1048
 FT CARBOHYD 1056 1056
 FT CARBOHYD 1084 1084
 FT CARBOHYD 1117 1117
 FT CARBOHYD 1133 1133
 FT CARBOHYD 1140 1140
 FT CARBOHYD 1148 1148
 FT CARBOHYD 1208 1208
 FT CARBOHYD 1233 1233
 FT CARBOHYD 1237 1237
 FT CARBOHYD 1332 1332
 FT CONFLICT 607 609
 FT CONFLICT 881 881
 FT SEQUENCE 1398 AA; 154474 MW; 355D873A27D56552 CRC64;
 Query Match 9.7%; Score 299.5; DB 1; Length 1398;
 Best Local Similarity 22.9%; Pred. No. 3.9e-14;
 Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;
 QY 96 LEETROXLEXTAKILID-----YIPDIATVVEY-----GDVXSKXXXT 135

Db 77 LEEAKTELEKLAELIDENRVLMMLVTKPERKVELNTSSLEKAMLRREKLSPEY 136
 QY 136 -HVESVEPLPYXIDDPOLFT-----KGASXVKKAKALDTKONKEV 176
 Db 137 KDKTKREP-----SLEPKRYNSTWYNALQFOEGCYDSSGVVAVLDTGVDPNHPFLSI 191
 QY 177 QLRGIEIXIAQXXXNDVXYITAKPEY-KVMN-----DVARGI----- 212
 Db 192 PFDGRKRIEMKDFDEGEFDTSFSPSKVNGTLINTTFQVASGLTNESTGLMEYVK 251
 QY 213 -----VKADVAQSSYGL-----GGOQIYAVADT 236
 Db 252 TVYVSNVTIGNTTSANGIYHFGLLPERKFDLNFDDQEDFYVLLVNSTGNDYDAAYDT 311
 QY 237 GLDTGRND-----SMHEAFRGKITALYALGRTNAN-----DTNGHGTHVAG 279
 Db 312 DLDYFTDEVPLAQYNTVYDVAVFSYTGPLNYVLAIEDPNCETAFVFGDHGCHTHVAG 371
 QY 280 SVLGNQXTN-----KGMAPQANLVFQSIMDSXG 307
 Db 372 TVAGYDSNNDDAMDWLSMYSGEWEVFSRLGWDYNTVTTDTGVGVAPGAQIMAIPLVRS-D 430
 QY 308 GLGLPLSNLQTLPSQAXSAGARIHTNSMGAVNGAYT--TDSRN--VDDYVRKNDMTILF 363
 Db 431 GRGSMWDLIEGM-TYAATGADVISMSLGG--NAPYLDGTDESYAVDELTEKYGVFEVI 487
 QY 364 AAGNENPNNGTISAGTANATVGCATEMLRSPFSYAD-----NTN 405
 Db 488 AAGNENPNNGTISAGTANATVGCATEMLRSPFSYAD-----NTN 405
 QY 406 HVAQFSSRGPTKGRIPDMVAPGTXTLSARSSLAPODSSFWANHDSKYVMGCTSMAPPI 465
 Db 546 RIAPFSSRRPRIDGELKRPVAVAPGYISLSLMMIGADP-----MGTSMAPPH 595
 QY 466 VAGNVAQLEHNFVKNRGTPKPSLKAALAGAA-----DXGLGTPNGCNGCKWGTLDK 519
 Db 596 VSGVVALLLSG-AKAEGLYNNPDIIRKYLESGATMLEGDPYQGYTELDOHGLVNVYK 654
 QY 520 SLNVAVYNSSKXLSQKATYFTATAGKPLISLVSAPASTTA---SVTLVNDLDLV 576
 Db 655 SWEI-----LKAINGTLPLIVDHMDKSIDPAEYIGVYINGILAR 696
 QY 577 ITAPN-----GTYXYGN---DEXXPXXNMW-----DG---RNVE----- 605
 Db 697 NSIPDIVEMHIXYGDTEKRFETIATPEWIKPFVSGSVILENNTFEVLRKYDVEGLEP 756
 QY 606 -----NVFINXQSGTYTTEVOAYNVPGQXFS 634
 Db 757 GLYVGRILIDPT--TPVIEDEILNTIVIPKEKT 788
 RESULT 4
 SUBV_BACSU STANDARD: PRT: 806 AA.
 AC P29141;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Minor extracellular protease Vpr precursor (EC 3.4.21.-).
 GN VPR OR IPA-45R.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=1423;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 161-195.
 RX MEDLINE=92041574; PubMed=1938892;
 RA Sloma A., Ruto G.A. Jr., Theriault K.A., Dwyer M., Wilson S.W.,
 RA Perc J.;
 RT "Cloning and characterization of the gene for an additional
 RT extracellular serine protease of Bacillus subtilis.";
 RL J. Bacteriol. 173:6889-6895(1991).
 RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN-168:
RX MEDLINE-95020537: PubMed-7934828:
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hillo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Prescan E., Sanaana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees."
RL MOL. Microbiol. 10:371-384(1993).
RP SEQUENCE FROM N.A.
RC STRAIN-168:
RX MEDLINE-98044033: PubMed-9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bartori M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Bourlier L., Brans A., Brann M., Brignell S.C., Bron S.,
RA Brouillet S., Bursch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.R., Codan J.U., Conneton I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmert P.T.,
RA Ertlan K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hillbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudea B., Park S.H.,
RA Paro V., Pohl T.M., Portetle D., Portolick S., Prescott A.M.,
RA Prescan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone C.,
RA Sekituchi J., Sekowska A., Seror S.J., Serrier P., Shin B.S., Soldo B.,
RA Sorochin A., Taccout E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoli A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassartol A.,
RA Viari A., Wambler K., Wedler E., Wedler H., Wetzinger T.,
RA Winters P., Wipal A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC - FUNCTION: NOT REQUIRED FOR GROWTH OR SPORULATION.
CC - SUBCELLULAR LOCATION: secreted.
CC - PTM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M76590; AAA22881.1; -
CC EMBL: X73124; CA51601.1; -
CC EMBL: Z59123; CAB15835.1; -
CC PIR: A41341; A41341.
CC HSSP: P00782; 2SPT.
CC MEROPS: S08.00A; -
CC Subtilisin: BG10591; VPR.
CC InterPro: IPR003137; PA.
CC Pfam: PF00082; Peptidase_S8; 2.
CC Pfam: PF02225; PA; 1.
CC PRINTS: PR00723; SUBTILISIN.
CC PROSITE: PS00136; SUBTILASE ASP; 1.
CC PROSITE: PS00137; SUBTILASE HIS; 1.
CC PROSITE: PS00138; SUBTILASE_SER; 1.

KW Hydrolase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 28
FT PROPEP 29 160
FT CHAIN 161 806
FT ACT_SITE 169 189
FT ACT_SITE 233 233
FT ACT_SITE 534 534
SQ SEQUENCE 806 AA; 85608 MW; F984E3BF0869DDDC CMC64;
Query Match 9.28; Score 286; DB 1; Length 806;
Best Local Similarity 22.58; Pred No. 1.8e-13;
Matches 132; Conservative 60; Mismatches 188; Indels 206; Gaps 19;
QY 155 TRKASLYVAXXALDTRKQNKVEQ--LGI-----EXIQXXXXNDYVYITAKREYKVN 206
DB 87 TRKKNKAIK--AVKNGKNREYEVGSFSGMKAPANEIKRLAAYDVAAVPPNTYITDN 144
QY 207 DVARGIVKADVAQS-----SYLYGOGIVAVADTGLDTR----- 242
DB 145 MKDKDVTISEDVAVSQMDSDAPYIGANDAMDLDYTGKGIKVAIIDTGEVYHHPDLKKNFG 204
QY 243 -----NDSSMHEAFRGKITALVALGRTNNANDNGHGVASVYLGXGXTNKGMAP 293
DB 205 QYKGYDFVNDYDPKEPTG-----DPRGEATDGHGVAGVFAANG--TIKGVAP 252
QY 294 QANLVFQSIIMDSXGGLGIPENLOTLESOAXSAGARIHTNSGAAVNGAYTTDSRNVDDY 353
DB 253 DATLLAVRYLIGFG--SGTTEENVLAGVERAVQDGADVMNLISGNSLNPPDVAATSTAL-DW 309
QY 354 VRKNDMTIFFAAGNEXPMGSGTISAPGAKNAITVGATE-----NLPRPSGVS----- 400
DB 310 AMEGGVAVTSGNNGSPRGWTVGSPTGSRSAISVGTQLPLMEYAVVTGTSASAKVGVN 369
QY 401 -----ADNI----- 404
DB 370 KEDDVVALNKKEVELVEAGIEADPFGRDLTKVAVVVRGSIADVDAADAKKAAGIAG 429
QY 405 -----NHVAQF 410
DB 430 VYNNLSGEIEANVPQSVPTIKLSLEDGKLVSAKACETTYTKRLTVSRALGEQVADF 489
QY 411 SSRGPKRD-GRIKPDVMAPGTYILSARSSLPADSSFWAMHDSKVAVMGSTMAATPIVAGN 469
DB 490 SSRGPMVDIMIKIPDISAPGVNIIVSTPTHPD-----HPYGYSGKGTBMASPHIAGA 543
QY 470 VAOLREFFVKNRGITPKPSL-LKAAALIAGAA-----DXGLGYPNGOGMGRTYLDKSLNV 523
DB 544 VAVTKQ-----AKPKWSVEQIKAAIMNTAVTLKDSDEYVPHNAGCAGSARI--MNA 593
QY 524 AYVNESSXLSTSQKATYXFTATPAGKPKISLVMSDAPASTASVTL 569
DB 594 --IKADSLVSPGSYGVGFLEKENEKNETFTTEMOSSIRKSYTL 637
RESULT 5
SUBT_BACS9
ID SUBT_BACS9 STANDARD: PRT; 420 AA.
AC P28842;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin precursor (EC 3.4.21.62).
GN SUB1.
OS Bacillus sp. (strain TA39).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92256481: PubMed-1581352;
RX Nartinx E., Davail S., Feller G.,
RT "Nucleotide and derived amino acid sequence of the subtilisin from
RT the antarctic psychrotroph Bacillus TA39."
RL Biochim. Biophys. Acta 1131:111-113(1992).

```

CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pI. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: STILL ACTIVE AT TEMPERATURES CLOSE TO 0 DEGREES
CC CELSIUS, IT HAS A MARKED HEAT LABILITY.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOROIATION. AND MANY MUTATIONS WHICH BLOCK SPOROIATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOROIATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62369; CAA44227.1; -
DR PIR: S23407; S23407.
DR HSSP: Q09405; 1MPT.
DR MEROPS: S08.0PA; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF000082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Sporulation; Serine protease; zymogen; signal.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 111 POTENTIAL.
FT CHAIN 112 420 SUBTILISIN.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 182 182 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 360 360 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 420 AA; 44086 MW; A64F121BD32826EC CRC64;

Query Match 8.98; Score 274.5; DB 1; Length 420;
Best Local Similarity 28.38; Pred. No. 5.4e-13;
Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

```

```

RESULT 6
PROA_XANCP
ID PROA_XANCP STANDARD: PRT: 580 AA.
AC P23314;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Extracellular protease precursor (EC 3.4.21.-).
GN XCC0851.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90251253; PubMed=2187155;
RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Danleis M.J.;
RT "A multipurpose broad host range cloning vector and its use to
RT characterise an extracellular protease gene of Xanthomonas campestris
RT pathovar campestris."
RL Mol. Gen. Genet. 220:433-440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reznach F.C., Farah C.S., Furian L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarelle G., Cannavan F., Cardozo J., Chambergo F., Clapina L.P.,
RA Claretelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.R., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51635; CAA35962.1; -
DR EMBL: AE012184; AAM40166.1; -
DR PIR: S11890; S11890.
DR HSSP: P00782; 25BT.
DR MEROPS: S08.0PA; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF000082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Serine protease; zymogen; signal.
FT SIGNAL 1 32 POTENTIAL.
FT PROPEP 33 2136 POTENTIAL.
FT CHAIN 2137 580 EXTRACELLULAR PROTEASE.
FT ACT_SITE 177 177 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 237 237 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 409 409 CHARGE RELAY SYSTEM (BY SIMILARITY).

```

Query Match 8.8%; Score 273.5; DB 1; Length 580;
 Best Local Similarity 22.7%; Pred. No. 9.8e-13;
 Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;

```

OY 1 MRKKVFLSYLSAALITLVALXNPSACXAXXFPDLKGIQTTDXGFSKQXQXGAFAF 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6 LKRTGSLTILGASALTSLLLMPAFAG-----EYIDGLATQHQKLTIVYKXGSL 60
OY 61 LVESENKTKLKGKLETPPANNNKHIXQFNPILEETKQXLEXGAKIILYIDYAYI 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 ASPS-----ALTSILRTAARAVPA-----KACKALGLNSVRLKGLVLY-- 100
OY 121 VERGGDVXKXXXXIEHVESEVEPEYLFXYIIDQFTKASXLYKAAALDTRKXNKEYOLRG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 101 -----RADRALDRABAEETLMQLADPNVOSVEV----- 129
OY 181 IETIAQXXXSNDVXYITANP-EYKVMNDVARGIVKA--DYAOSSXYGLYGGQIVAVADTG 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 130 -----DQILHATLTLPNDTRLESEOMAFGTNAGLNIINPAMDKATGSGTVAVADTG 179
OY 238 L---DTGRNDSMHEAFRGKITALYALGRNNAND-----TNGHGT 275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 180 ITSHADLNANILAGDYFISDATTARDGNGRDSNADGDMYANNECGAGIPAASSWHGT 239
OY 276 HVAGSY--LGNGXTN-KGNAPQANLYFGSINDXSG-----LGLIPENL 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 HVATTVAAVNNNTGVAGTAIGAKKVPYRVLGKCGSLSDIADAIYASGTVSGITPANA 299
OY 317 Q--TLFQASAGARHTMSQAAVNGAYTDSRNVDDYRKNDMLILPAAGNEPNGGT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 300 NPAYIVMSIGGGSGSTTMON-ALNGAVISGT-----TVVAAGADSNVSG 346
OY 375 ISAGCTKKAITYGATEN--LPPSGSADININVAOFSSRGPTKGRKIPVMAPGTYI 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 347 -SLPANCANIAVAATTSIGAKAASISNGGTGT-----DVSAPEGSSI 386
OY 433 LSAHSS--LAPDSSFWANHDSRYAVGCTSMATPIVAGNVAQLREHFVKNRGITPK--PS 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 387 LSTINSGTTPPGS-----ASYASYNGTSMASPHVAGVALVQS--VAPTLATPAAVET 437
OY 489 LLK--AALLIAGADKGLG-----PNCNGCMGRVTLDKSLNVAAYNESSX 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 438 LKNTATATLPGACGCGAGIYNADAAVTAALINGSGGGGNTLNGTPTVIGL----- 493
OY 532 LSTOKATYFTATAGKPLKISLVMSDAPASTASVTL--VMDLVI---TAPNGTXY 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 494 AATGAELNITITVPAG-----SGTILVTITSGSGSDADLIYRAGSAFTDSAY 539
OY 586 VGNDFXXXKXNNDGNHNNENFIMXPPSGITYTEVOAYN 625
DB 540 TCRPIRS-----GNAETCTITAP-SGIYVRLKAYS 569
  
```

RESULT 7
 WPA_BACSU STANDARD: PRT: 894 AA.
 AC PS4423: 006726;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell wall-associated protease precursor (PC 3.4.21.-) [Contains: cell
 wall-associated polypeptides CWBP23 and CWBP52].
 GN WPA.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.

RC STRAIN=168;
 RX MEDLINE=97158234; PubMed=9004506;
 RA Margot P., Karamata D.;
 RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
 RT growth, encodes a cell-wall-associated protease.";
 RL Microbiology 142:3437-3444(1996).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98015415; PubMed=9353931;
 RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
 RT "Sequencing of regions downstream of addA (98 degrees) and citC (289
 RT degrees) in Bacillus subtilis.";
 RL Microbiology 143:3305-3308(1997).
 RP [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Allison G.,
 RA Azevedo V., Bertier M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brigneau S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Etienne K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Golligly E.J., Grand G.,
 RA Giuseppe G., Guy B.J., Haga K., Hachech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchini M., Klein C.,
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Laryevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Meisel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Mostel D., Nakai S., Noack M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portecelle D., Potwolk S., Prescott A.M.,
 RA Prascan E., Puig C., Roca E., Roche B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield B.,
 RA Seiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Solde B.,
 RA Sotokuni M., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi A., Tamakoshi A., Tanaka T., Terstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wamboldt R., Wedler E., Medler H., Wetzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: NOT YET KNOWN. COULD BE INVOLVED IN PROTEOLYTICAN
 CC DEGRADATION BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
 CC -!- SUBCELLULAR LOCATION: CELL-WALL BOUND.
 CC -!- PTM: PROCESSED INTO CWBP23 AND CWBP52.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U58681; AAC25926.1; -;
 DR EMBL: U09476; CAAT0641.1; -;
 DR EMBL: 299109; CAB12917.1; -;
 DR HSSP: Q45670; IDBI.
 DR MEROPS: S08.004; -;
 DR Subtilist: BG11846; WPA.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISTIN.
 DR PROSITE: PS00136; SUBTILINASE_ASP; FALSE_NEG.
 DR PROSITE: PS00137; SUBTILINASE_HIS; 1.

```

DR PROSITE: PS00138; SUBTILASE_SER; 1.
KM Hydrolase; Serine protease; Cell wall; Zymogen; Signal;
KW Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 894
FT PROPEP 32 894
FT CHAIN 32 894
FT PROPEP 32 894
FT ACT_SITE 466 466
FT ACT_SITE 497 497
FT ACT_SITE 650 650
FT ACT_SITE 650 650
FT CONFLICT 9 9
FT CONFLICT 14 14
SQ SEQUENCE 894 AA; 96487 MW; 0F67C35B5F8DBC CRC64;

Query Match
Best Local Similarity 23.9%; Score 262; DB 1; Length 894;
Matches 134; Conservative 92; Mismatches 211; Indels 124; Gaps 28;

OY 49 FSKQXQTGAALF--LVESENVKLGKLETPA-----NNKLHIXQFN---G 93
DB 245 FQDVSENGASSKYTEKQKAIINRYDKALQSPSEFLKEIKKQADRLMKLOQKTAG 304
OY 94 PLLE-----TKQLXETGKILIDYIPDYIY-----EYEG-----DYSXXX 133
DB 305 ALTFENNTAAKSEVOPT--KVFVKDKNSLSSVHNEMKGFSAQSKKDIDSNVKKAKKL 362
OY 134 IEHVESVEPYLPYXYIDPOLFTFGASXLVKAALDPTQXNKREVLQIGEXIAQ-XXXND 192
DB 363 FENLVSFE--LPKDEKONAVYASAKRVKSAATLSKMSNVEF---AEPOEYKSLAND 416
OY 193 VXYITAKPEPKYKAMDVANGIV-ADVA-OSSYGLYGQGO---YVAVDPTGLDGRNDSS 246
DB 417 IYQ---PYQWPLKNNGENGVKNADYKEPANTLLSKRLNDTLAIVDPTGVDSFLAD-- 471
OY 247 MHEAFRGKITALYA---IGRTNNANDTNGHGFHVGSVLC---NGCTNKGMAPQANLVQ 300
DB 472 ---LKGKRTDLGHNFVGRNNANADDOGHGVAGILIAQSDNGSMTGNAAKATIPV 527
OY 301 SIMDSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAANGAVYTTDSRVNDDYVRKNDMT 360
DB 528 KYLDASG--SGPTQDLAIGIKYAADKGAKVINLSIG---GGYSRLLEALKYAADKNVL 581
OY 361 ILFAANGNEKPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHAQFSSRGPTKGR 420
DB 582 IAAAGSGNGEN--ALSTYPASSKYVMSVGAT-----NRMDWTADFMSYNGKGL-- 625
OY 421 IKPDVMAPTGXTILSARSSILAPDSSFWANHDSKYAYMGTSMATPIYAGVADLREHFVN 480
DB 626 ---DISAPGSDI-----PSLVPNGN-----VTYMSGTSMATPYAAAAAGLLFAPNPKL 670
OY 481 RGITRPSILKAA--LIAGAADXG-----LGYP-----NGNQMGHVTLLDKSLN 522
DB 671 KTEVEDMLKKTADDISFESVDGGEELDYDGDPIETIKPTGVVHSHSGYGLNWKAYS 730
OY 523 VA---YVNESXLTSGKAT 539
DB 731 AADQLKVNKLESTQTAVRGS 751

RESULT 8
ELYA_BACSP STANDARD: PRT; 378 AA.
AC P20724;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline elastase yab precursor (EC 3.4.21.-).
GN ALE.
OS Bacillus sp. (strain yab).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE-89359181; Pubmed-2670913;
RA Kaneko R., Koyama N., Tsai Y.-C., Juang R.-Y., Yoda K., Yamasaki M.;
RT "Molecular cloning of the structural gene for alkaline elastase yab,
a new subtilisin produced by an alkalophilic Bacillus strain.";
RL J. Bacteriol. 171:5232-5236(1989).
RN [2]
RP SEQUENCE OF 111-164.
RA Tsai Y.-C., Lin Y.-T., Li Y.-F., Yamasaki M., Tamura G.;
RT "Characterization of an alkaline elastase from alkalophilic Bacillus
yab-".
RL Biochim. Biophys. Acta 883:439-447(1986).
CC - FUNCTION: DIGEST ELASTIN EFFICIENTLY, HAS A SUBSTRATE PREFERENCE
FOR ALA IN P1 POSITION.
CC - SUBCELLULAR LOCATION: Secreted.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M28537; AA87324.1; -.
DR PIR; A33973; A33973.
DR HSSP; Q94405; 1MPT.
DR MEROPS; S08.0PA; -.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_SER; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 27
FT CHAIN 28 110
FT PROPEP 111 378
FT ACT_SITE 141 141
FT ACT_SITE 171 171
FT ACT_SITE 324 324
SQ SEQUENCE 378 AA; 38793 MW; 5A8F8C0C62687D CRC64;

Query Match
Best Local Similarity 8.3%; Score 257.5; DB 1; Length 378;
Matches 115; Conservative 50; Mismatches 167; Indels 99; Gaps 20;

OY 73 LKKLETVPANNKLHX-QFNGPILETRKQXLE--XTGAKILIDYIPDYIYETGDXVS 129
DB 1 MNKMGKIVAGTALIVSAFSSSIAQAAEEAKEKYLGKDEQVNSQ--FYDOIIGDEYS 58
OY 130 XXXXIEHVESVEPYLPYXYIDPOLFTFGASXLVKAALDPTQXNKREVLQIGEXIAQXXX 189
DB 59 ISSQAEDEVI--DLIHEFPIVLSVELDPEVDLLEDDPA----- 97
OY 190 SNDVXYITAKPEPKYKAMDVANGI--VKADVAQSSYGLYGQGOIVAVDPTGLDGRNDSSM 247
DB 98 ---IAYIEEDAETVMQYVPMGINRVQAPIAQR--GFTGTGVRAVLDPTGI-----SN 146
OY 248 HEAFRGKITALYALGRTNNANDTNGHGFHVGSV--LGNGXTNKGMAPQANLVFOSIM-- 303
DB 147 HADLRIRGASFPVGP--NISDGNHGQVAGTIALNLSIGVLAVPVDLYGKVLGA 205
OY 304 DSXGGLGGLPSNLQTLFSGQASAGARIHTNSWGAANGAVYTTDSRVNDDYVRKNDMTILF 363
DB 206 SGGSGISGIAQGLQ---WAANNGMHIANMSIGSSAGSSTMGQAVN--QATASGLVVA 258
OY 364 AAGNEXPNCGTISAPGTAKNAITVGATE--NLRPSFGSYADNINHAQFSSRGPTKGR 421
DB 259 ASGNS--GAGNGVFPARRANMAAVGATIDONNRKATFESQYAGL----- 299
OY 422 KPDVMAPTGXTILSARSSILAPDSSFWANHDSKYAYMGTSMATPIYAGVNA----- 471

```

```

Db 300 -DIVAPGV-----QSTVPNG-----YASFGTSMATPHVAGVAALVQKNPSWS 345
Oy 472 -QUREHFKVKN 480
Db 346 NVOIRNH-LKN 355

RESULT 9
SUBN_BACNA STANDARD: PRT: 381 AA.
ID SUBN_BACNA
AC P35835:
DF 01-JUN-1994 (Rel. 29, Created)
DF 01-JUN-1994 (Rel. 29, Last sequence update)
DF 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin NAT precursor (EC 3.4.21.62).
OS APRN.
GN Bacillus subtilis var. natto.
OC Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=86029;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC-2;
RX MEDLINE=93113095; PubMed=1369081;
RA Nakamura T., Yamagata Y., Ichishima E.;
RT "Nucleotide sequence of the subtilisin NAT gene, aprN, of Bacillus
RT subtilis (natto).";
RL Biosci. Biotechnol. Biochem. 56:1869-1871(1992).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS WITH BROAD SPECIFICITY
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pI. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOULATION, AND MANY MUTATIONS WHICH BLOCK SPOULATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOULATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D25319; BAA04989.1; -
DR EMBL: S51909; AAC60424.1; -
DR PIR: JH0778; JH0778.
DR HSSP: P07518; 1MEF.
DR MEROPS: S08.044.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASIN; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Signal.
FT SIGNAL 1 23
FT PROPEP 24 106
FT CHAIN 107 381
FT ACT_SITE 138 138
FT ACT_SITE 170 170
FT ACT_SITE 327 327
SQ SEQUENCE 381 AA: 39507 MW: DAED4B16ED1BA092 CRC64;

Query Match 8.2%; Score 255.5; DB 1; Length 381;
Best Local Similarity 25.9%; Pred. No. 1.2e-11;
Matches 124; Conservative 46; Mismatches 160; Indels 149; Gaps 19;
Oy 1 MRKKYFSLVLAAILSTVALXNPSAGXARXEDDFKQITTTDXGFSXQXTGAAP 60

```

```

Db 1 MRKKYFSLVLAAILSTVALXNPSAGXARXEDDFKQITTTDXGFSXQXTGAAP 60
Oy 61 LVESENKVLKNG--LXKKLETVPANNKLIHXPNGLILEETKQXLEXTGAKILDYIPDA 118
Db 52 SAKKDVISEKGKQKQKQKRYNA-----AAATDEKAAVELKQKPSVA 95
Oy 119 YIYEYEGDVYXXXXXIEVESVEPYLPXYXIDPQLFTKGASLVKAXLADKQKNREYOL 178
Db 96 YVE-----DHAHEAOSV-PY-----GISQ-IKAPL----- 122
Oy 179 RGLEIXAQQXXNDVXYITTAPEKYNANDVARGIKYADVAOSSTGLGGQGIYAVADTGL 238
Db 123 -----HSGQITGSNKKYAVDSGI 141
Oy 239 DTGRDSSMHEAFRGKITAYALGRTNANDTNGHGTAVGSV--LGNGXTNQMAPPQAN 296
Db 142 DSSHPLNV---RG--GASFPSETNPYODGSSHGTAVAGTIALNNSIGLVAPSPAS 195
Oy 297 LVFOSIMDSXGGLGGLPSNLOTLEFSQAXSAGARHTNSGAAVNGAVTTSRNVDYVRK 356
Db 196 LVAVKVLDSGT--SGQYSWIINGIEMAINMDVYIMSLGCPGT--STALKTVVDKAVS 250
Oy 357 NDMTILFAGNEXPNCG--GTISAPGTAKNAITVGA--TENVLPSPGYSADININVAQFS 412
Db 251 SGIVVAAAAGNCGSSGSTVGYPAKPTSTIANGAVNSNQASFSVSGSEL----- 302
Oy 413 RGPTRDGRKRPDYVAPGTXILSARSSLAAPSSPMANHDSKYAVKMGCTSMATPIYAGVA 471
Db 303 -----DYMAFGVSI-----QSTLPQGT-----YGVNGTSMATPHVAGAA 338

RESULT 10
AQLI_THEAO STANDARD: PRT: 513 AA.
ID AQLI_THEAO
AC P08594;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aqualysin I precursor (EC 3.4.21.-).
GN PSTI.
OS Thermus aquaticus.
OC Bacteria: Thermus/Delnococcus group: Delnococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_Taxid=271;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 15-23.
RC STRAIN=YT1.
RX MEDLINE=90216674; PubMed=2182621;
RA Terada I., Kwon S.-T., Miyata T.;
RT "Unique precursor structure of an extracellular protease, aqualysin
RT I, with NH2- and COOH-terminal pro-sequences and its processing in
RT Escherichia coli.";
RL J. Biol. Chem. 265:6576-6581(1990).
RN [2]
RP SEQUENCE OF 75-442 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=YT1.
RX MEDLINE=88225062; PubMed=2386255;
RA Kwon S.-T., Terada I., Matsuzawa H., Ohta T.;
RT "Nucleotide sequence of the gene for aqualysin I (a thermophilic
RT alkaline serine protease) of Thermus aquaticus Yr-1 and
RT characteristics of the deduced primary structure of the enzyme.";
RL Eur. J. Biochem. 173:491-497(1988).
RN [3]
RP SEQUENCE OF 128-170.
RX MEDLINE=88151937; PubMed=3162211;
RA Matsuzawa H., Tokugawa K., Hamachi M., Mizoguchi M., Taguchi H.,
RA Terada I., Kwon S.-T., Ohta T.;
RT "Purification and characterization of aqualysin I (a thermophilic
RT alkaline serine protease) produced by Thermus aquaticus Yr-1.";
RL Eur. J. Biochem. 171:441-447(1988).
CC -1- FUNCTION: AQUALYSIN I IS A THERMOPHILIC ALKALINE SERINE PROTEASE.
CC THE OPTIMAL TEMPERATURE FOR ITS CASEINOLYTIC ACTIVITY IS 80

```


[8]
RN ACTIVE SITE.
RX MEDLINE=69104413; PubMed=5249818;
RA Markland F.S., Shaw E., Smith E.L.;
RT "Identification of histidine 64 in the active site of subtilisin";
RL Proc. Natl. Acad. Sci. U.S.A. 61:1440-1447(1968).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE.
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in P1. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted
CC -1- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
CC Alcalase by Novozymes
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOROULATION AND MANY MUTATIONS WHICH BLOCK SPOROULATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOROULATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 58.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02496; AAB05345.1; -.
DR PIR: A00970; SUBSN.
DR PDB: 1S01; 15-OCT-90.
DR PDB: 1S02; 15-JAN-92.
DR PDB: 1SBH; 07-DEC-95.
DR PDB: 1SB1; 07-DEC-95.
DR PDB: 1SBN; 31-JAN-94.
DR PDB: 1SRT; 31-MAY-84.
DR PDB: 2SRT; 31-OCT-93.
DR PDB: 1STB; 31-OCT-93.
DR PDB: 2STC; 15-APR-93.
DR PDB: 3STC; 31-JAN-94.
DR PDB: 5STC; 31-JAN-94.
DR PDB: 2SNT; 15-JAN-93.
DR PDB: 1SEB; 15-OCT-95.
DR PDB: 2ST1; 15-JUL-91.
DR PDB: 1ST2; 15-JUL-91.
DR PDB: 1SUA; 14-JAN-98.
DR PDB: 1SUB; 31-JAN-94.
DR PDB: 1SUC; 31-JAN-94.
DR PDB: 1SUD; 31-JAN-94.
DR PDB: 1SUE; 14-OCT-98.
DR PDB: 1SUP; 14-NOV-95.
DR PDB: 1AK9; 12-NOV-97.
DR PDB: 1AUG; 31-DEC-97.
DR PDB: 1AON; 14-JAN-96.
DR PDB: 1YJA; 11-JUL-96.
DR PDB: 1YJB; 11-JUL-96.
DR PDB: 1YJC; 11-JUL-96.
DR PDB: 1A20; 29-APR-98.
DR MEROPS: S08.034.
DR InterPro: IPR00209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Sporulation; Serine protease; Zymogen; Signal;
KW 3D-structure.
FT SIGNAL 1 32 POTENTIAL.
FT PROPEP 33 107
FT CHAIN 108 382 SUBTILISIN BPN'.
FT ACT_SITE 139 139 CHARGE RELAY SYSTEM.
FT ACT_SITE 171 171 CHARGE RELAY SYSTEM.
FT ACT_SITE 328 328 CHARGE RELAY SYSTEM.

FT HELIX 113 117
FT TURN 118 119
FT HELIX 120 125
FT TURN 126 127
FT TURN 131 132
FT STRAND 134 139
FT TURN 144 145
FT TURN 147 148
FT STRAND 151 156
FT TURN 159 160
FT TURN 164 165
FT HELIX 171 180
FT STRAND 183 194
FT STRAND 196 201
FT TURN 205 206
FT HELIX 211 223
FT TURN 224 225
FT STRAND 228 231
FT STRAND 235 235
FT HELIX 240 251
FT TURN 252 253
FT STRAND 255 259
FT TURN 267 268
FT STRAND 274 274
FT TURN 275 277
FT TURN 279 280
FT STRAND 282 287
FT TURN 289 290
FT STRAND 293 293
FT TURN 295 296
FT TURN 301 302
FT STRAND 305 308
FT TURN 312 316
FT STRAND 317 319
FT STRAND 320 324
FT HELIX 327 344
FT TURN 346 347
FT TURN 350 359
FT HELIX 360 360
FT STRAND 362 362
FT STRAND 367 370
FT TURN 371 372
FT STRAND 374 374
FT HELIX 377 381
SQ SEQUENCE 382 AA; 39181 MW; ED987DAFA37B8335 CRC64;

Query Match 8.2%; Score 253; DB 1; Length 382;
Best local similarity 24.0%; Pred. No. 1.8e-11;
Matches 120; Conservative 43; Mismatches 148; Indels 188; Gaps 17;

QY 1 MRKKVFLSVLSAAIIISTVALNMPASGAKRFDLPKGIOTTTXGFEKXOTGAAP 60
DB 1 MRKKVWISLPLALALITTFMFGSTSAQAAG-----KNGEKKIYVGFQTMSTMAA 54

QY 61 LVSENVKTLKGLKKLETPVANKKLHIQFNGPILUEFKOXLEXTGAKIIDYTPYAT 120
DB 55 -----KKKVVISEKGVQKQFK--YDAAASATLNKAVKELKQPSVAV 98

QY 121 VEYEGVXSXXXXIEH-----ESVEPYLPXYXIDPLFTKGASLYVAKALDTKQXNKE 175
DB 99 EE-----DHVAHVAQSV-PY-----GVSQ-IKAPAL----- 123

QY 176 VOLRGIEIXIAQXXSNDVXYITANPEYKVMNDVARGIVADVAQSYGLQGQIVAVAD 235
DB 124 -----HSQGYTGSNVKVAVID 139

QY 236 TGLDTRGNDSSMHEAFRGKITVALYALGRTNNANDTGHGTIVAGSV--LGNGXTNKGMAP 293
DB 140 SGID-----SSHPLKAVAGASVWPSETNPFDNNSHGTIVAGTVAALNNSIGVLGVAP 193

QY 294 QANLVQSTIMDSXGLGLPSNLQTLFSSQXXSAGARIHTNSMGAAVNCVATYTTDSRAND-- 351
DB 194 SASLVAVKVLGADG-----SGQSYWILNCEIEMALANNNDVI 229

```

QY 352 -----DYRKNDMTLLFAAGNEXPNG--GTISAPGTAKNAITVGATEN 392
D 230 NMSLGGPSSAALKAAYKAVAGVVAAGNNECTSSSSSTGYGPKYPSYLAAG--- 286
QY 393 LRPSFGSADININHAQSSSRGPKDKIKPDVWAPGXYILSARSLAPDSSFMANHRSK 452
D 287 -----VDSNNORASFSVGP-----ELDVNAPGVSIQSTLPG-----NK 320
QY 453 YAYMGTSMATPIVAGNVA 471
D 321 YGAYNGTSMASPHVAGAA 339

RESULT 12
SUBT_BACSA
ID SUBT_BACSA STANDARD: PRT: 381 AA.
AC P00783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin amylosacchariticus precursor (EC 3.4.21.62).
GN APP.
OS Bacillus subtilis var. amylosacchariticus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1483;
RN 1;
RP MEDLINE=89008194; PubMed=3139650;
RA Yoshimoto T., Oyama H., Honda T., Töne H., Takeshita T.,
RA Kamlyama T., Tsuru D.;
RT "Cloning and expression of subtilisin amylosacchariticus gene.";
RL J. Biochem. 103:1060-1065(1988).
RN 12;
RP PARTIAL SEQUENCE.
RA MEDLINE=72266687; PubMed=4560201;
RA Markland F.S., Kurihara M., Smith E.L.;
RT "Subtilisin Amylosacchariticus. II. Isolation and sequence of the
RT tryptic and cyanogen bromide peptides.";
RL J. Biol. Chem. 247:5602-5618(1972).
RN 13;
RP SEQUENCE OF 107-381.
RX MEDLINE=72266688; PubMed=5055784;
RA Kurihara M., Markland F.S., Smith E.L.;
RT "Subtilisin Amylosacchariticus. 3. Isolation and sequence of the
RT chymotryptic peptides and the complete amino acid sequence.";
RL J. Biol. Chem. 247:5619-5631(1972).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pI. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPORE germination, AND MANY MUTATIONS WHICH BLOCK SPORE germination AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPORE germination.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D00264; BAA00186.1; -
DR PIR, A00971; SUBSS.
DR PIR, A41448; A41448.
DR HSSP, P04189; ISCJ.
DR HSSP, S08_042; -
DR InterPro, IPR000209; Peptidase_S8.

```

```

DR Pfam, PF00082; Peptidase_S8; 1.
DR PRINTS, PR00723; SUBTILISIN.
DR PROSITE, PS00136; SUBTILISIN_ASP; 1.
DR PROSITE, PS00137; SUBTILISIN_HIS; 1.
DR PROSITE, PS00138; SUBTILISIN_SER; 1.
KM Hydrolyase; Sporulation; Serine protease; Zymogen; signal.
FT SIGNAL 1
FT PROPEP 31 106
FT CHAIN 107 381
FT ACT_SITE 138 138
FT ACT_SITE 170 170
FT ACT_SITE 327 327
FT CONFLICT 191 191
FT CONFLICT 365 365
SQ SEQUENCE 381 AA; 39467 MW; 2251BADE22B4824F CRC64;

Query Match 8.1%; Score 251.5; DB 1; Length 381;
Best Local Similarity 26.0%; Pred. No. 2.3e-11;
Matches 124; Conservative 49; Mismatches 159; Indels 145; Gaps 19;

QY 1 MRKKVFLSVLAAILSTVALXNPSAGXARFDFKGIQTIDXXGSKOXGTGAFAF 60
D 1 MRKKVFLSVLAAILSTVALXNPSAGXARFDFKGIQTIDXXGSKOXGTGAFAF 60
QY 61 LVESENVKLXG-LKKLETPANNKLIXQFNPILEETKQXLEXTGAKILDYIPDYA 118
D 61 LVESENVKLXG-LKKLETPANNKLIXQFNPILEETKQXLEXTGAKILDYIPDYA 118
QY 119 YIVEEGVYXKXKXIEHVESEVETPLPYXITDPLQFTGASXLKXKALDTRQKNKEVQL 178
D 119 YIVEEGVYXKXKXIEHVESEVETPLPYXITDPLQFTGASXLKXKALDTRQKNKEVQL 178
QY 96 YVEE-----DHIAHEYAQSV-PY-----GISQ-IKAPAL----- 122
D 96 YVEE-----DHIAHEYAQSV-PY-----GISQ-IKAPAL----- 122
QY 179 RGIEIXAQXXXXNDVXYITAKPEYKVMNDVARGIVKADVAOSYGLYQOGIVAVADTGL 238
D 179 RGIEIXAQXXXXNDVXYITAKPEYKVMNDVARGIVKADVAOSYGLYQOGIVAVADTGL 238
QY 123 -----HSQCYGTSNKKAVVIDSGI 141
D 123 -----HSQCYGTSNKKAVVIDSGI 141
QY 239 DTGRDSSMHEAFRGKITALYALGRTNANDTNGHTHVAGSV--LGNGXTNKGAPQAN 296
D 239 DTGRDSSMHEAFRGKITALYALGRTNANDTNGHTHVAGSV--LGNGXTNKGAPQAN 296
QY 142 DSHPLDNV-----RG--CASFPVSETPNYQDSSSHGTVAAGTIALNNSICVLAVPSAS 195
D 142 DSHPLDNV-----RG--CASFPVSETPNYQDSSSHGTVAAGTIALNNSICVLAVPSAS 195
QY 297 LVFOSIMDSXGGLGLPNSLQTLFQSAKSAGARITHNMGCAVAVGAYTTDSRNDVYRK 356
D 297 LVFOSIMDSXGGLGLPNSLQTLFQSAKSAGARITHNMGCAVAVGAYTTDSRNDVYRK 356
QY 196 LYAVKVIDSTG--SGQYSWIINGIEMALSNMMDYINMSLGP--SGSTAKTAYVDKAVS 250
D 196 LYAVKVIDSTG--SGQYSWIINGIEMALSNMMDYINMSLGP--SGSTAKTAYVDKAVS 250
QY 357 NDMTILFAAGNEXPNG--GTISAPGTAKNAITVGATENLRPSFGSADININHAQSSRG 414
D 357 NDMTILFAAGNEXPNG--GTISAPGTAKNAITVGATENLRPSFGSADININHAQSSRG 414
QY 251 SGIVVAALAGNMGSSSTGYGPKYPSYLAAG-----VSSNQRASFSFSG 299
D 251 SGIVVAALAGNMGSSSTGYGPKYPSYLAAG-----VSSNQRASFSFSG 299
QY 415 PTKDGRIPDVWAPGTYLARSLSLAPDSSFMANHRSKAYVMGTSMATPIVAGNVA 471
D 415 PTKDGRIPDVWAPGTYLARSLSLAPDSSFMANHRSKAYVMGTSMATPIVAGNVA 471
QY 300 S-----ELDVNAPGVSI-----QSTLPGT-----YGAYNGTSMATPHVAGAA 338
D 300 S-----ELDVNAPGVSI-----QSTLPGT-----YGAYNGTSMATPHVAGAA 338

RESULT 13
ELVA_BACAO
ID ELVA_BACAO STANDARD: PRT: 380 AA.
AC P27693;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus alcalophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1445;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=PB92;
RX MEDLINE=91282483; PubMed=2059048;
RA van der Laan J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
RA Quax W.J.;
RT "Cloning, characterization, and multiple chromosomal integration of a
RT Bacillus alkaline protease gene.";
RL Appl. Environ. Microbiol. 57:901-909(1991).

```

```

RN [2]
X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC STRAIN-PB92.
RA MEDLINE-9230330; PubMed-1518788;
RA van der Laan J.C., Teplyakov A.V., Kelders H., Kalk K.H., Misset O.,
RA Mulleers L.J.M., Dijkstra B.W.;
RA "Crystal structure of the high-alkaline serine protease PB92 from
RA Bacillus alcalophilus.";
RA Protein Eng. 5:405-411(1992).
RN [3]
X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RA MEDLINE-93078250; PubMed-144775;
RA Sobek H., Hecht H.-J., Aehle W., Schomburg D.;
RA "X-ray structure determination and comparison of two crystal forms of
RA a variant (psu15arg) of the alkaline protease from Bacillus
RA J. Mol. Biol. 228:108-117(1992).
RN [4]
STRUCTURE BY NMR OF 112-380.
RC STRAIN-PB92.
RA MEDLINE-9277237; PubMed-9115441;
RA Martin J.R., Mulder F.A., Karim-Nejad Y., van der Zwan J.,
RA Martini M., Schipper D., Boelens R.;
RA "The solution structure of serine protease PB92 from Bacillus
RA alcalophilus presents a rigid fold with a flexible substrate-binding
RA site.";
RA Structure 5:521-532(1997).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: M65098; AAA2212.1;
DR EMBL: A13738; CAA01128.1;
DR PIR: A49778; A49778.
DR PDB: 1AH2; 1S-APR-98.
DR MEROPS: S08.038;
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR HydroLase: Serine protease; zymogen; Signal; 3D-structure.
FT SIGNAL 1
FT PROPEP 28 112
FT CHAIN 113 380 ALKALINE PROTEASE.
FT ACT_SITE 143 143 CHARGE RELAY SYSTEM.
FT ACT_SITE 173 173 CHARGE RELAY SYSTEM.
FT ACT_SITE 326 326 CHARGE RELAY SYSTEM.
SQ SEQUENCE 380 AA; 38853 MW; 539EA72771B6682C CRC64;
Query Match 8.1%; Score 250.5; DB 1; Length 380;
Best Local Similarity 24.4%; Pred. No. 2,8e-11;
Matches 12; Conservative 57; Mismatches 148; Indels 177; Gaps 22;
OY 4 KKFVLSVAALSTVALXNPSAGXAFEDLDFKGIQTITDDXGFSKXQXGAAFLVE 63
DB 2 KKPGLKIVASTALLISVAFSSIASMAEAEKRY-----LIGNEQ----- 42
OY 64 SENVKLKGKLETPVANNKLHXQFNGLIEETKQXLEXTGAKILIDYIVAYVEY 123
DB 43 -----EAVSEFEQVEANDVAI-----LSEEEVEIE-----LHEF 75
OY 124 EG-DVXSGXXXIHEVSEVPEPLPYXIXIDPOLFTKGASGLVKAKALDTQXNKVEGLRTE 182
DB 76 ETIVLVSELPEDVDALF-----LDPA----- 98

```

```

OY 183 XIAQXXXNDVYXITAKPEYKM-NDVARGIVKADY-AQSSYGLYGQGIIVAVDTGLDT 240
DB 99 -----ISIEDEAEVTTMAQSPWMCISRYQAPAAHNGLTGSGVKAIVDTGIST 148
OY 241 GRNDSMHPAFKGTITVALYALGRTNNANDPNGHGTIVACSV--LNGGTNNKGMAPQANLV 298
DB 149 -----HPDLNIRGASFPVGP--STQDNGHGTIVAGTIAALNNISIGVGVAPNAELV 200
OY 299 FQSIDMSXGGLGSLPSNLTQTLFQASASGARITHNSMGA-----AVNGAYTTDSRN 349
DB 201 AVYVLASG--SGSVSSTIAQGLEMAKNGNHVANNLSGSPSATLLEQAVNSA---TSRG 255
OY 350 VDDYVKNMTILFPAAGDEXNGCTISAPGTAKNATTGATE--NLRSFSGYADNINHV 407
DB 256 V-----LVVAASGNS--GAGSISTPARYANMAVATQNNRASFSGTAGL--- 301
OY 408 AQPSSKQPTKDRIRKPDYMAPQTYILSARSSLAPDSSEFVANHDSKYAYMGTSNAPPIVA 467
DB 302 -----DLRHFVKN 480
OY 468 GNYA-----QLEHFVKN 480
DB 334 GAALVKGKQKPSMSVQIRNH-LKN 357
RESULT 14
ID ELYA_BACCS STANDARD; PRT; 380 AA.
AC P41362;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus clausii.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=9680;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-221 / ATCC 21522 / JCM 9139 / DSM 2512;
RX MEDLINE-93043753; PubMed-1368952;
RA Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,
RA Aono R., Horikoshi K.;
RT "Molecular cloning, nucleotide sequence, and expression of the
RT structural gene for alkaline serine protease from alkaliphilic
RT Bacillus sp. 221.";
RL Biosci. Biotechnol. Biochem. 56:1455-1460(1992).
RN [2]
RP SEQUENCE OF 112-129.
RC STRAIN-221 / ATCC 21522 / JCM 9139 / DSM 2512;
RA Horikoshi K.;
RL (in) Horikoshi K. (eds.);
RL Microorganisms in alkaline environments, pp.187-194, VCH,
RL Weinheim (1991).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
EMBL: S48754; AAC60420.1;
EMBL: D13157; BAA02442.1;
EMBL: A26817; CAA01836.1;
EMBL: A22550; CAA01611.1;
HSSP: P29600; IGCI.
MEROPS: S08.103;
InterPro: IPR000209; Peptidase_S8.
Pfam: PF00082; Peptidase_S8; 1.

```

```

DR PRINTS: PRO0723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
KW Hydrolyase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 27
FT PROPEP 28 111
FT CHAIN 112 380
FT ACT_SITE 143 143
FT ACT_SITE 173 173
FT ACT_SITE 326 326
SO SEQUENCE 380 AA; 38826 MW; 5F73ABC68D586831 CRC64;

Query Match      8.1%; Score 250.5; DB 1; Length 380;
Best Local Similarity 24.4%; Pred. No. 2.8e-11;
Matches 123; Conservative 57; Mismatches 148; Indels 177; Gaps 22;

OY 4 KKVFLSVLSAAILSTVALXNPSAGAXRFXDLDFKGIOTTTDXKSPSOXOTGAAPLVE 63
DB 2 KKPLGKIYASTALLISVAFSSISASAEAEKEKY-----LIGFNEQ----- 42
OY 64 SENVLXGKGLKLTETVANKKLIHXONGPILETKOXLETKGAKIIDYIPDAYIYEX 123
DB 43 -----EAVSEFEVQVANDQVAT-----LSEEEVEITE-----LHFE 75
OY 124 EG-DYKXKXKXKXIEHVESEPEYLPXYXIDPOLFTKGASXLVKAXALDTRQXNKVEQLRGIE 182
DB 76 ETIPLVSLSEPEDVDALD-----LDPA----- 98
OY 183 XIAQKXKXNDVXYITAKEPEYKX-NDVARGIYKADY-AOSSYGLYGOGIYAVADTGLDT 240
DB 99 -----ISYIEDAEVTTTMAQSVPMGSRVQAPAHNKGILSGSGKVAVLDLTGIST 148
OY 241 GRNDSMEHAEFGKITALYALGRTNANDTNGCHGVHVASV--LGNGXTNKGMAPQANLY 298
DB 149 -----HPDLIRGGAFFYFGEF-STODGNCHGVHAGTIALANNSIGVLGAFASAEIY 200
OY 299 FOSIMDSXGGLGGLPSNLQTFESQXKSAGARIHTNSMGA-----AVNGAYTTDSRN 349
DB 201 AVKVLGASG--SGSVSTIAQGLEFMAGNNGMHVANAISLSPSPSATLEQAVNSA---TSRG 255
OY 350 VDDYRKNDMTLFFAAGXENPGCFISAPGTAKNAITVGAPF--NLRSFGSYADININHV 407
DB 256 V-----LVVAASGNS--GAGSISTPARYANAAVAGATDOONNRASFQYAGL--- 301
OY 408 AOFSSRGPTKGRIRKPDYKMGTXILASRSSIAPDSFWANHDSKYAYMGTSMAPTIVA 467
DB 302 -----DIVAPGVNVQSYTPG-----STYASLNGTSMATPHVA 333
OY 468 GNVA-----QLRHFVKV 480
DB 334 GAALVYKOKNPSWMSNVQIRNH-LKN 357

RESULT 15
SOBT_BACLI
ID SUBT_BACLI STANDARD; PRT; 379 AA.
AC P00780;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Subtilisin Carlsberg precursor (PC 3.4.21.62).
GN APR.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 6816;
RX MEDLINE=86093688; PubMed=3001653;
RA Jacobs M., Eliasson M., Uhlen M., Flock J.-I.;
RT "Cloning, sequencing and expression of subtilisin Carlsberg from
RT Bacillus licheniformis.";
```

```

RL Nucleic Acids Res. 13:8913-8926(1985).
RN [2]
RP SEQUENCE OF 106-379.
RX MEDLINE=68234702; PubMed=4967581;
RA Smith E.L., Delange R.J., Evans W.H., Landon M., Martland F.S.;
RT "Subtilisin Carlsberg. V. The complete sequence; comparison with
RT subtilisin BPN'; evolutionary relationships.";
RL J. Biol. Chem. 243:2184-2191(1968).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF VARIANT WITH SELENOCYST-325.
RX MEDLINE=93291170; PubMed=8512925;
RA Syed R., Wu Z.P., Hoyle J.M., Hillvert D.;
RT "Crystal structure of selenosubtilisin at 2.0-A resolution.";
RL Biochemistry 32:6157-6164(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 106-379.
RX MEDLINE=98087517; PubMed=9425066;
RA Stoll V.S., Eger B.T., Hynes R.C., Martichonok V., Jones J.B.,
RA Pal E.F.;
RT "Differences in binding modes of enantiomers of L-acetamido boronic
RT acid based protease inhibitors: crystal structures of gamma-
RT chymotrypsin and subtilisin Carlsberg complexes.";
RL Biochemistry 37:451-462(1998).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IF CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pl. Hydrolyzes peptide amides.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
CC Alcalase by Novozymes.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOULATION, AND MANY MUTATIONS WHICH BLOCK SPOULATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOULATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X03341; CAB56500.1; -.
DR PIR: A00968; SUBSC.
DR PIR: A24111; SUBSCL.
DR PDB: 1CSE; 15-OCT-89.
DR PDB: 1SCA; 31-JAN-94.
DR PDB: 1SCB; 31-JAN-94.
DR PDB: 1SCD; 31-JAN-94.
DR PDB: 1SBC; 15-JAN-95.
DR PDB: 2SEC; 15-JAN-95.
DR PDB: 1SCN; 31-AUG-94.
DR PDB: 1AF4; 16-JUN-97.
DR PDB: 1SET; 31-OCT-93.
DR PDB: 1VSB; 18-MAR-98.
DR PDB: 3VSB; 25-MAR-98.
DR PDB: 1AVT; 01-APR-98.
DR PDB: 1AV7; 01-APR-98.
DR PDB: 1BE6; 14-OCT-98.
DR PDB: 1BE8; 13-JAN-99.
DR PDB: 1BFK; 18-NOV-98.
DR PDB: 1BFU; 18-NOV-98.
DR MEROPS: S08.001; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PRO0723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Signal;
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:35:29 ; Search time 60.4527 Seconds
(without alignments)
217.969 Million cell updates/sec

Title: US-09-920-954-1

Sequence: 1 MRKKVFLSVLSAAAILSTV.....EVQAYNPVPGPQXESLAIVN 639

Scoring table: BLOSUM62

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 10

Listing first 45 summaries

Database ::

```

1: SPSEMBL_21:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mmc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.prodent:*
13: sp.virus:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.archaeal:*
17: sp.archaea:*

```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2155	69.5	434	2	Q9A0R3	Q9A0R3 bacillus sp
2	2155	69.5	434	2	Q93UV9	Q93UV9 bacillus sp
3	2082	67.1	444	2	Q9A0R0	Q9A0R0 bacillus sp
4	1952.5	63.0	433	2	Q9A0R1	Q9A0R1 bacillus sp
5	1948.5	62.8	433	2	Q9A0R2	Q9A0R2 bacillus sp
6	1941.5	62.6	433	2	Q9A0R4	Q9A0R4 bacillus sp
7	193	15.9	1835	2	Q9A0R2	Q9A0R2 bacillus sp
8	478	15.4	1702	5	Q9E1W1	Q9E1W1 dictyosteli
9	412	13.3	361	16	Q9B9N7	Q9B9N7 dictyosteli
10	408	13.2	654	17	Q8BUC2	Q8BUC2 thermoaer
11	381	12.6	1233	16	Q9B2A4	Q9B2A4 pyrococcus
12	368.5	11.9	1253	16	Q9B2A4	Q9B2A4 streptomyc
13	356	11.3	1102	2	Q9B2A4	Q9B2A4 streptomyc
14	350.5	11.3	444	16	Q9A0R7	P95684 streptomyc
15	337	10.9	1245	16	Q9A0R5	Q9A0R7 bacillus ha
16	336.5	10.9	412	2	Q9A0R6	Q9A0R5 streptomyc
						Q9A0R6 thermoaer

17	336.5	10.9	412	16	08RC68
18	321.5	10.4	1220	16	09LO00
19	319.5	10.3	442	16	031788
20	315				045464 bacillus su
21	310	10.0	824	2	045464
22	301.5	9.7	861	2	093635
23	295.5	9.5	1398	1	099P11
24	281.5	9.0	431	2	099S16
25	280	9.1	444	2	054327
26	279.5	9.0	1155	1	099P01
27	276.5	8.9	719	2	045681
28	270.5	8.7	499	16	09KEM1
29	270	8.7	397	2	P97097
30	270	8.7	621	2	09P466
31	270	8.7	621	2	053401
32	268.5	8.7	379	2	066153
33	260.5	8.4	715	2	P70765
34	257.5	8.3	1345	1	054437
35	255	8.2	403	2	045463
36	255	8.2	692	2	09EXK0
37	253	8.2	1098	16	09L118
38	250.5	8.1	629	2	093R68
39	250.5	8.1	374	2	099P43
40	250.5	8.1	374	2	099P42
41	250.5	8.1	378	2	045466
42	250	8.1	601	2	046540
43	250	8.1	379	2	045300
44	249.5	8.0	352	2	045522
45	249.5	8.0	374	2	099P94
			757	16	09K666
					091668 thermotomax
					091060 streptococcus
					031788 bacillus su
					045464 bacillus su
					093635 thermococcu
					099111 pyrococcus
					093616 bacillus sp
					054327 bacillus sp
					099p01 uncultured
					045681 bacillus su
					09kemi bacillus su
					p97097 bacillus sp
					09f466 alteromonas
					053401 alteromonas
					066153 bacillus sp
					p70765 alteromonas
					p70765 staphylothe
					054543 bacillus sp
					09exk0 bacillus ps
					091128 streptococcus
					093r68 alteromonas
					09f943 bacillus ll
					09f942 bacillus ll
					045466 bacillus sp
					046540 bacteroides
					045300 bacillus ll
					045322 bacillus sp
					09f941 bacillus ll
					09k666 bacillus su

ALIGNMENTS

RESULT 1			
09A0R3			
ID	09A0R3	PRELIMINARY;	PRT: 434 AA.
AC	09A0R3:		
DT	01-JUN-2001 (TREMBlrel. 17, Created)		
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)		
DE	Protease (Fragment).		
GN	PROA.		
OS	Bacillus sp. 9860.		
OC	Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Bacillaceae: Bacillus.		
OX	NCBI_TaxID=133778;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=9860;		
RC	MEDLINE=2058675; PubMed=11118284;		
RA	Seki K., Okuda M., Hatada T., Kobayashi T., Ito S., Takami H.,		
RA	Horikoshi K.;		
RT	"Novel oxidatively stable subtilisin-like serine proteases from		
RT	alkaliphilic Bacillus spp.: enzymatic properties, sequences, and		
RT	evolutionary relationships."		
RL	Biochem. Biophys. Res. Commun. 279:313-319(2000).		
DR	EMBL: AB046403; BAB2166.1; -.		
DR	HSSP: P00782; ISUP.		
DR	InterPro: IPR00209; Peptidase_S8.		
DR	Pfam: PF00082; Peptidase_S8; 2.		
DR	PRINTS: PR00723; SUBTILISIN.		
DR	PROSITE: PS00137; SUBTILASIN_HIS; UNKNOWN_1.		
DR	PROSITE: PS00138; SUBTILASE_SER; 1.		
FT	NON_TER	1	
FT	NON_TER	1	
FT	NON_TER	434	
FT	NON_TER	434	
SEQUENCE	434 AA; 45311 MW; APC9F78B8143527E CRC64;		
Query Match		69.5%;	Score 2155; DB 2; Length 434;
Best Local Similarity		96.3%;	Pred. No. 2.le-122;
Matches 418:	Conservative	0;	Mismatches 16; Indels 0; Gaps 0;

```
|||||
Db 1 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 60
Oy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVPOSINDSXGGLGCLPSNLQTLFSQAXS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVPOSINDSXGGLGCLPSNLQTLFSQAXS 120
Oy 326 AGARHTMSWGAANVGATTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
Db 121 AGARHTMSWGAANVGATTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
Oy 386 TVGATENLRPSPFGSYADININHAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 445
Db 181 TVGATENLRPSPFGSYADININHAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 240
Oy 446 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADXGIGY 505
Db 241 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADVGLGY 300
Oy 506 PNGNGMGWRVTLDKSLNVAAYNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTTA 565
Db 301 PNGNGMGWRVTLDKSLNVAAYNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Oy 566 SVTLVNDLDVLTAPNGTXYVGNDFXPPXXNMGRNVENFVIXXPOSTYTTIEVOAYN 625
Db 361 SVTLVNDLDVLTAPNGTXYVGNDFXPPXXNMGRNVENFVIXXPOSTYTTIEVOAYN 420
Oy 626 VPVGPQXFSLAIVN 639
Db 421 VPVGPQXFSLAIVN 434
```

RESULT 2

```
093UV9 PRELIMINARY: PRT: 434 AA.
ID 093UV9:
AC 093UV9:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Protease (Fragment).
GN PROE.
OS Bacillus sp. KSM-KP43.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=109322;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KP43.
RA Itoh S., Saeki K.;
RT "new protease.";
RL Submitted (NCV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB051423; BAB55674.1; -
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 434
SQ SEQUENCE 434 AA: 45302 MM: 0B08418191853CA3 CRC64;

Query Match 69.5%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 2,le-122;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
Oy 206 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 265
Db 1 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 60
Oy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVPOSINDSXGGLGCLPSNLQTLFSQAXS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVPOSINDSXGGLGCLPSNLQTLFSQAXS 120
Oy 326 AGARHTMSWGAANVGATTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
```

```
|||||
Db 121 AGARHTMSWGAANVGATTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
Oy 386 TVGATENLRPSPFGSYADININHAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 445
Db 181 TVGATENLRPSPFGSYADININHAOPSSRGPTKDGRIKPDVMAPGTXILSARSLAPDSSF 240
Oy 446 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADXGIGY 505
Db 241 WANHDSKYAYMGTSMAPIVAGNVAOLREHFVKNRGITTPKPSLKALILAGADIGIGY 300
Oy 506 PNGNGMGWRVTLDKSLNVAAYNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTTA 565
Db 301 PNGNGMGWRVTLDKSLNVAAYNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Oy 566 SVTLVNDLDVLTAPNGTXYVGNDFXPPXXNMGRNVENFVIXXPOSTYTTIEVOAYN 625
Db 361 SVTLVNDLDVLTAPNGTXYVGNDFXPPXXNMGRNVENFVIXXPOSTYTTIEVOAYN 420
Oy 626 VPVGPQXFSLAIVN 639
Db 421 VPVGPQXFSLAIVN 434
```

RESULT 3

```
09AOR0 PRELIMINARY: PRT: 434 AA.
ID 09AOR0:
AC 09AOR0:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROE.
OS Bacillus sp. NV1.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=133781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NV1.
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL: AB046406; BAB21269.1; -.
DR HSSP: P00782; 1SOP.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 2.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 434
SQ SEQUENCE 434 AA: 45294 MM: 83517EDDB74125D2 CRC64;
```

```
Query Match 67.1%; Score 2082; DB 2; Length 434;
Best Local Similarity 91.5%; Pred. No. 5,3e-118;
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

Oy 206 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 265
Db 1 NDVARGIKADVAOSSYGLYGOGQIVAAVDGLDGRNDSSMHEAFRGKITALVALGRTN 60
Oy 266 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVPOSINDSXGGLGCLPSNLQTLFSQAXS 325
Db 61 NANDNGHGTTHVAGSVLNGXTNKGMAPQANLVPOSINDSXGGLGCLPSNLQTLFSQAXS 120
Oy 326 AGARHTMSWGAANVGATTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
Db 121 AGARHTMSWGAANVGATTTDSRNVDYVRKNDMAVLFAAGNEXPNGGTISAPGTAKNAI 180
```



```
QY 386 TVGATENLRPSFGSYADNINNHVAQFSSRGPTKGRKIPDVMAFGTXLLSARSSLAPDSSF 445
      |||
DB 181 TVGATENLRPSFGSYADNINNHVAQFSSRGPTKGRKIPDVMAFGTXLLSARSSLAPDSSF 240
QY 446 WANIDSKYAVMGSTSMATPIVAGNVAOLREHFVKNRGITPSPILKALILAGAADGIGY 505
      |||
DB 241 WANIDSKYAVMGSTSMATPIVAGNVAOLREHFVKNRGITPSPILKALILAGAADGIGY 300
QY 506 PNGNGMGRTYTLKSLNVAAYVNESSXLTSTOKATYXFTATAGPKLISLWMDAPASTTA 565
      |||
DB 301 PNGNGMGRTYTLKSLNVAAYVNESSXLTSTOKATYXFTATAGPKLISLWMDAPASTTA 360
QY 566 SVTLVNDLVLITPAFGTXVGNDFXPPXXXNMGRRNVEVFINKPGSGTYTIEVOAYN 625
      |||
DB 361 SVTLVNDLVLITPAFGTXVGNDFXPPXXXNMGRRNVEVFINKPGSGTYTIEVOAYN 420
QY 626 VPGPQXFSLAIVN 639
      |||
DB 421 VPGPQXFSLAIVN 434
```

RESULT 4

```
Q9AOR1 PRELIMINARY; PRT: 433 AA.
ID Q9AOR1
AC Q9AOR1
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. SP521.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=133780;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS521;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046405; BAB21268.1; -.
DR HSSP; Q45670; IDBI.
DR InterPro; IPR000209; Peptidase-S8.
DR Pfam; PF00082; Peptidase_S8; 2.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 433
FT SEQUENCE 433 AA; 45576 MW; 9842DF18FE660DDC CRC64;
```

```
Query Match 63.0%; Score 1952.5; DB 2: Length 433;
Best Local Similarity 86.6%; Pred. No. 3.4e-110;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;
QY 206 NDVARGIVKADVAAQSSYGLYGOGIYAAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 265
      |||
DB 1 NDVARGIVKADVAAQSSYGLYGOGIYAAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 60
QY 266 NANTNGHGHVAGSVLNGXTNKGMAPOANLVFQSIIMDSXGGLGCLPSNLDTLFSQAXS 325
      |||
DB 61 NANTNGHGHVAGSVLNGXTNKGMAPOANLVFQSIIMDSXGGLGCLPSNLDTLFSQAXS 119
QY 326 AGARIHNSGAAVNGAYTTDSRNDVYVRKNDMTILFAAGNEXPNRGITSAAGTAKKAI 385
      |||
DB 120 AGARIHNSGAAVNGAYTTDSRNDVYVRKNDMTILFAAGNEXPNRGITSAAGTAKKAI 179
QY 386 TVGATENLRPSFGSYADNINNHVAQFSSRGPTKGRKIPDVMAFGTXLLSARSSLAPDSSF 445
```

```
DB 180 TVGATENYRRPSFSLADNPNHIAQFSSRGATRDGRKIPDVATAGTFTLSARSSLAPDSSF 239
      |||
QY 446 WANIDSKYAVMGSTSMATPIVAGNVAOLREHFVKNRGITPSPILKALILAGAADGIGY 505
      |||
DB 240 WANIDSKYAVMGSTSMATPIVAGNVAOLREHFVKNRGITPSPILKALILAGAADGIGY 299
QY 506 PNGNGMGRTYTLKSLNVAAYVNESSXLTSTOKATYXFTATAGPKLISLWMDAPASTTA 565
      |||
DB 301 PNGNGMGRTYTLKSLNVAAYVNESSXLTSTOKATYXFTATAGPKLISLWMDAPASTTA 359
QY 566 SVTLVNDLVLITPAFGTXVGNDFXPPXXXNMGRRNVEVFINKPGSGTYTIEVOAYN 625
      |||
DB 361 SVTLVNDLVLITPAFGTXVGNDFXPPXXXNMGRRNVEVFINKPGSGTYTIEVOAYN 419
QY 626 VPGPQXFSLAIVN 639
      |||
DB 420 VPGPQXFSLAIVN 433
```

RESULT 5

```
Q9AOR4 PRELIMINARY; PRT: 433 AA.
ID Q9AOR4
AC Q9AOR4
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=127869;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D6;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046402; BAB21265.1; -.
DR HSSP; Q45670; IDBI.
DR InterPro; IPR000209; Peptidase-S8.
DR Pfam; PF00082; Peptidase_S8; 2.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 433
FT SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;
```

```
Query Match 62.8%; Score 1948.5; DB 2: Length 433;
Best Local Similarity 86.4%; Pred. No. 6e-110;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;
QY 206 NDVARGIVKADVAAQSSYGLYGOGIYAAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 265
      |||
DB 1 NDVARGIVKADVAAQSSYGLYGOGIYAAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 60
QY 266 NANTNGHGHVAGSVLNGXTNKGMAPOANLVFQSIIMDSXGGLGCLPSNLDTLFSQAXS 325
      |||
DB 61 NANTNGHGHVAGSVLNGXTNKGMAPOANLVFQSIIMDSXGGLGCLPSNLDTLFSQAXS 119
QY 326 AGARIHNSGAAVNGAYTTDSRNDVYVRKNDMTILFAAGNEXPNRGITSAAGTAKKAI 385
      |||
DB 120 AGARIHNSGAAVNGAYTTDSRNDVYVRKNDMTILFAAGNEXPNRGITSAAGTAKKAI 179
QY 386 TVGATENLRPSFGSYADNINNHVAQFSSRGPTKGRKIPDVMAFGTXLLSARSSLAPDSSF 445
      |||
DB 180 TVGATENLRPSFGSYADNINNHVAQFSSRGATRDGRKIPDVATAGTFTLSARSSLAPDSSF 239
```

```
QY 446 MANDSKYAYMGSTMAATPIVAGNVAOLREHFVKNRGITPKPSLILKALAGAADXGLGY 505
|||:|||||
Db 240 MANNNSKYAYMGSTMAATPIVAGNVAOLREHFITKNRGITPKPSLILKALAGATDVGLGY 299
PNNOGGGRVTLKSLNVAAYVNESSXLSSTOKATYXTATAGPKLISLWSDAPASTTA 565
QY 506 PNNOGGGRVTLKSLNVAAYVNESSXLSSTOKATYXTATAGPKLISLWSDAPASTTA 565
|||:|||||
Db 300 PNDGCGGRVTLKSLNVAAYVNEATALTGTGKATYSFOTAGPKLISLWTDAPGSTTA 359
SVTLVNDLDELVTAPNGCTYVGNDFXPPXNNMDGRNNVENVEFINXPOSGTYTIEVOAYN 625
QY 566 SVTLVNDLDELVTAPNGCTYVGNDFXPPXNNMDGRNNVENVEFINXPOSGTYTIEVOAYN 625
|||:|||||
Db 360 SYTLVNDLDELVTAPNGOKYVGNDFSPYDNNMDGRNNVENVEFINAPQSGTYTIEVOAYN 419
VVPVPOXFSLAIVN 639
QY 626 VVPVPOXFSLAIVN 639
|||:|||||
Db 420 VPSGPORFSLAIVH 433

RESULT 6
Q9AOR2 PRELIMINARY; PRT; 433 AA.
ID Q9AOR2;
AC Q9AOR2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROC.
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=133779;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL: AB046404; BAB21267.1; -.
DR HSSP: Q45670; IDBI.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PFO0082; Peptidase_S8; 2.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45587 MW; B81291A803C775AE CRC64;

Query Match 62.6%; Score 1941.5; DB 2; Length 433;
Best local similarity 86.2%; Pred. No. 1.6e-109;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 206 NDVARGIVKADVAQSSYGLYGQGIYAADVADTGLDGTGRNDSMHEAFRGKITALAYALGRTN 265
|||:|||||
Db 1 NDVARGIVKADVAQNNGYGLYGQGVAAVADTGLDGTGRNDSMHEAFRGKITALAYALGRTN 60
NANDTNGHGHVAGSVLGNXTKKGMAPOANLVFOSTMDXSGLGIGPSLQTLFSSQAXS 325
Db 61 NASDPNGHGHVAGSVLGN-ALNKGMAPOANLVFOSTMDSSGGLGPSLNLTLFSSQAWN 119
AGARIHTNSGAVNGAYTTDSRNVDYVRKNDMTILFAAGNEXPNGGTSAPGTAKNAI 385
Db 120 AGARIHTNSGAVNGAYTANSRQVDEYVANNMTVLFAAGNGPNSGTSAPGTAKNAI 179
TVCATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMABGTXYLSARSSLADPSSP 445
Db 180 TVGATENLRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTXYLSARSSLADPSSP 239
MAHNDKSYALMGSTMAATPIVAGNVAOLREHFVKNRGITPKPSLILKALAGAADXGLGY 505
```

```
Db 240 MANNNSKYAYMGSTMAATPIVAGNVAOLREHFITKNRGITPKPSLILKALAGATDVGLGY 299
PNNOGGGRVTLKSLNVAAYVNESSXLSSTOKATYXTATAGPKLISLWSDAPASTTA 565
QY 506 PNNOGGGRVTLKSLNVAAYVNESSXLSSTOKATYXTATAGPKLISLWSDAPASTTA 565
|||:|||||
Db 300 PNDGCGGRVTLKSLNVAAYVNEATALTGTGKATYSFOTAGPKLISLWTDAPGSTTA 359
SVTLVNDLDELVTAPNGCTYVGNDFXPPXNNMDGRNNVENVEFINXPOSGTYTIEVOAYN 625
QY 566 SVTLVNDLDELVTAPNGCTYVGNDFXPPXNNMDGRNNVENVEFINXPOSGTYTIEVOAYN 625
|||:|||||
Db 360 SYTLVNDLDELVTAPNGOKYVGNDFSPYDNNMDGRNNVENVEFINAPQSGTYTIEVOAYN 419
VVPVPOXFSLAIVN 639
QY 626 VVPVPOXFSLAIVN 639
|||:|||||
Db 420 VPSGPORFSLAIVH 433

RESULT 7
Q8T9W1 PRELIMINARY; PRT; 1825 AA.
ID Q8T9W1;
AC Q8T9W1;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Serine protease/ABC transporter TagD.
GN TAGD.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Anjard C., Loomis W.F.;
RT "Evolution of the ABC transporters of Dictyostelium.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF466309; AAL74253.1; -.
KW Protease.
SQ SEQUENCE 1825 AA; 202641 MW; E28160BC78613A3B CRC64;

Query Match 15.9%; Score 493; DB 5; Length 1825;
Best local similarity 24.2%; Pred. No. 2.3e-21;
Matches 192; Conservative 105; Mismatches 229; Indels 268; Gaps 31;

QY 87 HIXQFNGPILLET-----KQXLEXTG-----ANILDIYDPYAYIV---EYEGDVXSX 130
|||:|||||
Db 127 YIVQFKDHINDETRFQKFLINTDIVLDEQPYQSHIVNYIPHDSFLVLMNDEQSMLSS 186
XXXIIEHVESYEP-----YLRXXYXIDPQLEFTGASXL-----YK 163
Db 187 KEMVSWTGEFEPSNKTILHLYNNEKSGISGLPYIKLSDSTNSLIQWENTLNSILTSYNSKYA 246
AXALDTRQXNKEVQOLNGIEXIAQXXS-----NDVXYITAKPEYKVMNDY 208
Db 247 LTLINQKTL-KSLVYCNDESSQSGSLVSEKLYVQWISQESYNIETSEKFORANRL 305
ARCIV-----KADVAQSSYGLYGQGIYAADVADTGLDGTGR--NDS-----SM 247
QY 209 ARCIV-----KADVAQSSYGLYGQGIYAADVADTGLDGTGR--NDS-----SM 247
|||:|||||
Db 306 SPKAIETGKTDLVNNRIDIP-----LRGQGLISLADTGLDGHCFSDSNNDIPYNSV 360
HEAFRGKITALAYALGRTNANNDTNGHGHVAGSVLNGXTN-----KGMAPQANLVFO 300
Db 361 NLNHRKVYVYIGSL--HDNEDYVDGHTHVCAGAACAPEDSSLAISFSLADAKIAF 418
SI-MDSXGVLGIGPSLQTLFSSQAXSAGARIHTNSMGA-----AVNGAYTTDSRNVDYVR 355
Db 419 DLASDPSNNEPVPEDYPSQLYQPLVNAAGARVHDSMSLSIQGLSYSDASIDDFLY 478
KN-DMTILFPAGEXXPGCGTISAPGTAKNAITYGATENLRPS----- 397
QY 356 KN-DMTILFPAGEXXPGCGTISAPGTAKNAITYGATENLRPS----- 397
THPDFILRAAGNNEQYSSLSL-QATAKNVITVGABQTTHESTTDALDYSNFETVAKST 537
Db 479 THPDFILRAAGNNEQYSSLSL-QATAKNVITVGABQTTHESTTDALDYSNFETVAKST 537
GSYAD-----NINHVAQ 409
QY 398 -----GSYAD-----NINHVAQ 409
|||:|||||
Db 538 LNSLQSGFDDKYCTYTTAOCCTEYSTVKGISGCTCTYIKNSTASIFSSOPELYENNNIS 597
```

```

OY 410 FSSKPTKGRIRPDVMAFGYXILSARSSLA-----PDSEFWANDSKTAYVWG 458
DB 598 FSSKGPVHDLRDPDIAFGQYITARSNGANTTDOCGDGLPTNTALLSE-----SG 650
OY 459 TSMATPIVAGNVAOLREH-----FYXNRGITRPEPLKALILAGADUXGLGYP-- 506
DB 651 TSMATPLATAATTLIRQYLDVGYPTGSIYSNKLQPTGSLKALMINNAQOLNLTGFP 710
OY 507 -----NGNCGMGRVTLDKSLNVAYNES----- 530
DB 711 STNTNPSNAVFDPFAGANFVQMGSLRMEWL---YVESGKVPKPSRWVGIGELGRDK 767
OY 531 -----XLISTOKATYXFT-----ATAGKP-LKLSLWSDAPASTASVTLVND 574
DB 768 ASMKREYSLSTGONVSYCTYKPRSSSGNSGCIPIRYATVLTDPDPSYSGAKLNLVND 827
OY 575 LVIT-----APNGXYYG-NDFXPPXAMWDRNNVNF---INXQSGITTE 620
DB 828 LTMNTSEPIFYNSGSGSYNGTGLPLQ---DSINVEGILTYTPTNKSEISFRT 884
OY 621 VQATNVPGFQXS 634
DB 885 IAGTNPIDGPNFS 898

RESULT 8
O9GTN7 PRELIMINARY; PRT: 1702 AA.
AC O9GTN7
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE TagA.
GN TAGA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA "TagA, a putative serine protease/ABC transporter of Dictyostelium
RT that is expressed at the onset of development and is required for the
RL differentiation of a subpopulation of prespore cells."
DR EMBL; AF263455; AAG11416.1; -.
DR HSSP; P13569; INBD.
DR Interpro; IPR003593; AAA_Atpase.
DR Interpro; IPR001140; ABCtransporter.
DR Interpro; IPR003439; ABC_transporter.
DR Interpro; IPR002029; Peptidase_S8.
DR Pfam; PF00664; ABC_membrane_1.
DR Pfam; PF00005; ABC_tran_1.
DR PRINTS; PR00723; SUPRILISIN.
DR PRODOM; PD000006; ABC_transportr_1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; 1702 AA; 187103 MW; 4A67716303CE7131 CRC64;
SQ
SEQUENCE 1702 AA; 187103 MW; 4A67716303CE7131 CRC64;

Query Match 15.4%; Score 478; DB 5; Length 1702;
Best Local Similarity 21.9%; Pred. No. 1,7e-20;
Matches 167; Conservative 125; Mismatches 261; Indels 210; Gaps 26;
OY 75 KKLTVPANRKLIXQFNGPILEETKQXL-----EXTGAKIDYIPDAYIVVEGCG-- 126
DB 78 KKSINQSKSGSLFLVHLNGPIEKQVHNLIKQLDPLNGGEIHHYIPDMTYILSMISGDN 137
OY 127 -----VXSXXXIIEHVESVEPYLPYXITD----- 151
DB 138 NDNNNNKILINRKLKELPSIQWKLPLEPRKLSPLEFKQNGQFQGNQNDQLKATYYHE 197
OY 152 -----QLETKGASLVVAKXALDTRKQXNKE-----VQLR-----GIEXIAQXXSND 192

```

```

DB 198 NSNQSDNDNTITSESSLTVEKELISNNNNNNNVLITVNNKSKLSESTIYKIFRS 257
OY 193 VXYITAKPEYKYV-----NDVARGIVKADVAQS-----YGLYGQGIYVADPTGLT 240
DB 258 LVYIEPSSSKLIRKTPSKNKAHVSIOGSASTSPIMDVIGIKGDEIYVACADVTGDI 317
OY 241 GR--NDSSMHEAFRKTALYALGRTNNANDTNGHTVAGSVLNGXTN-----KGM 291
DB 318 NHCFEYDNPIGISTHKKIIS--YSSGNGDQIDELDGQTHVGTIIGSTTVDPSEVSEFG 376
OY 292 APOANLVFOSINDSGKGGAGLPNQTLFQOAXSAGARHTNSMGAANVAGATYDSRVD 351
DB 377 APNSKAFVLDLVGSGNSGLSIQSNLTALYQSTYDQANKVHCAMNSNIGPFTYGTGEMID 436
OY 352 DYARKN-DMTILEAGNEXPNG-----GISAPGARNAITVAGTEMLRPSGVSADIN 406
DB 437 RFQWDHEDFLVYASAGNVNFGNSYITLSQESTSKNSLVGSSN--QPS-STYLSIDY 493
OY 407 ----- 406
DB 494 WMDFIYNSIRSVCTQOOSIYITCSDVPTQTSYDIQTCCSNPILAKICCTEIQOQ 553
OY 407 -----VAOFFSKRPTKGRIRPDVMAFGYXILSARSSLAPOSSFWANH----- 449
DB 554 YQTNSTVYSEPIPSLFSGVGPTSDGRRLKPDLLAGPSIISR-SLSPSTI--NHCSPT 610
OY 450 ----DSKYAVMGTSMAPIVAGNVAOLRE-----HFVKNR-----GITPKPSLLKALI 495
DB 611 SGITSAIIMEGSSQAAAVATSAAVLVROYRRDGYFINKVSSVGFQPSASVKKALI 670
OY 496 AGAA--DKLQCPNGNOGCVTLDKSLNVAYNES-----XLISTOKATY 540
DB 671 NTASINVDSTLEY---SQGFNGIQLSKITTTNAGTSTLSDIPSTIEKADPIINGETNSY 727
OY 541 XFRATGKPLKSLVMSDAPASTASVTLVNDLYTLA-PNC--TXVYGNQ---FXXPX 594
DB 728 CFSLDKADIDITLVMTDPAGSPSLSTFLVANNLDLALLAFVDELSTIYSGNSERTIFKNTS 787
OY 595 XXMMDGRNNEVNFINKPQSGTYTIEVQAINVPGFQXESLAI 637
DB 788 QVLEFDLNNVEYIRIKDAPIGSYDKIFGTNIYIPNOSVSVYI 830

RESULT 9
O9RBJ2 PRELIMINARY; PRT: 561 AA.
AC O9RBJ2
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Subtilisin-like serine proteases.
GN APRE2 OR TTE0824.
OS Thermomicrobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermomicrobacteriales; Thermomicrobacteriaceae; Thermomicrobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-MB4T / JCM11007;
RA MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013049; AAM24081.1; -.
KW Protease; Complete proteome.
SQ
SEQUENCE 561 AA; 59968 MW; BA9C5C52F7083A18 CRC64;

Query Match 13.3%; Score 412; DB 16; Length 561;
Best Local Similarity 27.6%; Pred. No. 4.1e-17;
Matches 142; Conservative 74; Mismatches 177; Indels 122; Gaps 19;

```



```

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrett B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL391041; CAC01588.1; -.
DR HSSP: Q99405; IMPT.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF02225; PA.1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 1239 AA; 128505 MW; 8F5E9AC68EB1260A CRC64;

Query Match 12.6%; Score 391; DB 16; Length 1239;
Best Local Similarity 30.9%; Pred. No. 2e-15;
Matches 143; Conservative 51; Mismatches 187; Indels 82; Gaps 16;

OY 213 VKADVAQSSY-----GLYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALG 262
DB 219 VEADLDSTAOIGAFRAWAGGNTGGVEAVLDYGVAG-----HPDLADRIAAQSFV 272
OY 263 RTNNANDTNGHGTAVGSLVNGXTN---KGMAPQANVFQSIINDSGXGGGLPSMLQRT 318
DB 273 PBDNTDGRGHTHASTIACTGAASGKKGVAFCARLSICKVLDN--SGRQJSMTLAA 331
OY 319 LFSQAKSAGARLHTNSMGAANCAVYTTD--SRNVDYVFRNDWTIIFAGNEXPRGCTISA 377
DB 332 MEMAAVERIAKLVNMSLSGSGSDSPMSRAVDRLSAOTGALFVYAGN--CGEAGSIGA 390
OY 378 PCTAKNATVTCATENLRPSFGSYADNINHYAOPSSKGPDKRIKIPDVMAFGTYLSARS 437
DB 391 PGVAISALTVGA-----VDATDILAPFSSGCPVVDGALKPEITAGVGIILAA-- 437
OY 438 SLAPDSFMANHDSKYAVMGCTSMATPIYAGVNAQOLREHFVNKRGITPRKPSILKAAL-- 494
DB 438 ----NSSFAAGNGAYOSLSGTSMTAPHYAAGAAALL-----AAARPLSGSALKQV 484
OY 495 IGAADXGIGYPNGNCGMGRVTLDKSLN-----VAVNBSXLSLSOKATYFTATAG 547
DB 485 LASSSHRTPRDADFQAGSGRYDVAARAGVYASATAYAPGSSPGPYARLVTYNTTGA 544
OY 548 KPLKISLVMSDA-----PASTTASVTLVNDLVLITAPNGTYVGNDFXPX 593
DB 545 VTLEISVATITAPGVEFRLSASRYTVPAHGTAQVLTLLIDGS---GSAAGRAYSGQIL 601
OY 594 XXXXNDGRRNENVFINKXPSGTYTTEVO--AYNVEVGFQXFSL 635
DB 602 A----DARNVAHTAVSAGPVRAKRLTVHFKDQAGNEV--PGVFDL 639

RESULT 12
O99C06 PRELIMINARY; PRT: 1253 AA.
AC 099C06
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE Putative secreted peptidase.
GN SC0176 OR: SC8A11.04C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;

[1]
RN SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Cerdeno A.M., Parkhill J., Barrett B.G., Rajandream M.A.;
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Kadenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinsahl H., Hopwood D.A.;
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RN M01. Microbiol. 21:77-96(1996).
RN SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Penley S.D., Chatter K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrett B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL391041; CAC01576.1; -.
DR HSSP: Q99405; IMPT.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF02225; PA.1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 1253 AA; 130971 MW; AA69B417FEEDB89 CRC64;

Query Match 11.8%; Score 368.5; DB 16; Length 1233;
Best Local Similarity 28.0%; Pred. No. 4.7e-14;
Matches 142; Conservative 55; Mismatches 169; Indels 141; Gaps 20;

OY 205 MNDVARGIVKADVQSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALGRT 264
DB 219 LSDTYAOIGADV--WGGNTGGVEAVLDYGVAG-----HPDRGRITATASFVDP 270
OY 265 NNANDTNGHGTAVGSLVNGXTN---KGMAPQANVFQSIINDSGXGGGLPSMLQRT 318
DB 271 QDVYDRNNGHGTAVGSLVNGXTN---KGMAPQANVFQSIINDSGXGGGLPSMLQRT 318
OY 308 -----GLGGLPSMLQTLFSQAKSAGARLHTNSMGAANCAVYTTDSRNDVYR 355
DB 331 AVRDGHAKIVSMISGSDPTDGTDPLEAV-----NMLSAETGA----- 368
OY 366 KNDWTILFPAAGNENXPGTISAPGTAKNATVTCATENLRPSFGSYADNINHYAOPSSKRP 415
DB 369 ----LFLVAAGNSGPEAVYITGPRAADALITGVANG--PQK--VDQLADPSKRP 417
OY 416 -TKQRLKPRDVMAPGTYIISARSILAPDSFMANHDSKYAVMGCTSMATPIYAGVNAQOLR 474
DB 418 RYGDNAVKPDLTAPGVEGLAARSRYAPES-----EGAYOSLSGTSMTAPHYAAGAAALL 471
OY 475 EHFVKNKGITPRKPSLKAALILAGAADXGIGYPNGNCGMGRVTLDKSLNVAVYVNSXLSL 534
DB 472 LBNPDWTG-----QRLKELVGTAGTGRFSP--FDASGKV-----DVAAVRSVTLAS 519

```

QY 535 SQ-----KATY-----XFTATAGKPLKISLWSDA-----PAST 563
 DB 520 GDAFAQAHYPYFGQTVRRVRYTNSGPAVALDLSPALPEGLFTLSEAOVTPAIG 579
 QY 564 TASYTLVNDL-----LVITAPNGTXYVGNDFXPPXXNMNDRNNVENVTI--- 609
 DB 580 TASYGVITHLDAEDNGAYATRLVAGADGAVLA-----RTPVGNREGRATLATAKDH 635
 QY 610 -NXPQSGTYTIEVOAYNVFPGQFSL 635
 DB 636 HDKPLSGTILKDYERN--TAPKVIYSV 660

RESULT 13
 P95684 PRELIMINARY; PRT: 1102 AA.
 ID P95684
 AC P95684
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Subtilisin-like protease.
 OS Streptomyces albobroiseolus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=1887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S-3253;
 RX MEDLINE-97144528; PubMed-8990295;
 RA Suzuki M., Taguchi S., Yamada S., Kojima S., Momose H.;
 RT "A novel member of the subtilisin-like protease family from
 RT Streptomyces albobroiseolus.";
 RL J. Bacteriol. 179:430-438(1997).
 DR EMBL: D83672; BAA12040.1; -
 DR HSP: P00782; 258T.
 DR Interpro: IPR002860; GH_BNR.
 DR Interpro: IPR000209; Peptidase_S8.
 DR Pfam: PF02012; BNR; 2.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Protease.
 SQ SEQUENCE 1102 AA; 114128 MW; F9E4AD2590FE559E CRC64;

Query Match 11.5%; Score 356; DB 2; Length 1102;
 Best Local Similarity 30.5%; Pred. No. 2,3e-13;
 Matches 144; Conservative 55; Mismatches 185; Indels 88; Gaps 19;

QY 203 KYMNDYAR-----GIVADYAOQ-----SYGLYGQQLYAVADTGLDTRNDSSMH 248
 DB 178 RASAGVAVRWLDGVRKASLDTSVGQIGTPKAMAGYDGKVKIAYLDTCVD-----ATH 231
 QY 249 EAFRGKITLALVAGRTNNANDTNGHGTAVAGSVLGNX-----TNKGMADPANIWFQSIIND 304
 DB 232 PDLKGVYTSKNTSAPTTGVDYVGHGTAVASIAAGTGAOSKGTGYKAVAGAKILNKGVD 291
 QY 305 SXGGLGSLPSNIQTLFSQAXSAGARIHTNSKGAANVAGATTTDSRNDVYRK--NMTIL 362
 DB 292 DNG--FGDDSGILAGEMAAQCADIVNNSLG---GMDTPETDPLEAIVDKLSAEKGL 345
 QY 363 F--AAGNEXPNGGTSAPATKAKNAITVAGATENLRPSFGSYADININVAOFSSKRP--TKDG 419
 DB 346 FAIAAGNEPQ--STGSPSADSALTIVA-----VDDKDLADFPSSITGPRLGDG 392
 QY 420 RIKPDVMAAGTYILSARSSLADDSFMANHDSKAYAMGTSNATPIYAGNVQALREHFVK 479
 DB 393 AVKPDLTAGVDITASAKGNDIAKEVGEPAGYWTISCTSNMTPHYAGAAALIKQHPD 452
 QY 480 NNGITPSPBLKALAGAADXGLGTPNGNQGRTYLDKSLNVAIVNESSXLS----- 533

DB 453 -----WKYAEKLGALTASTKDG--KYTFEGQSGRVQVDKAITQVIAEPLSFGVQM 505
 QY 534 -----TSOKATYXFTATAGKPLKISLWSD-----APAS--TTASVTLVNDLDTVITA 579
 DB 506 PHADKPVTKKTLRLNGLGEDYTLKLTSTATGPKXGAAPAGFTTLCASLT-----IV 557
 QY 580 P-NGTYVGNDFXPPXXNMNDRNNVENVFINXPOS-----GYTTEVOAYNV 626
 DB 558 PANGTASVDVTDTRLGAAGVDGTYSAYVATAGQSGVFTAAVEREYESNV 609

RESULT 14
 Q9KBJ7 PRELIMINARY; PRT: 444 AA.
 ID Q9KBJ7
 AC Q9KBJ7
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Intracellular alkaline serine protease.
 GN APRX OR BH1930.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE-20512582; PubMed-11058132;
 RA Takami H., Nakasone K., Takaki Y., Meeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001513; BAB05649.1; -
 DR HSP: Q99405; 1MP.
 DR Interpro: IPR001128; Cytochrome_P450.
 DR Interpro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Protease; Complete proteome.
 SQ SEQUENCE 444 AA; 48916 MW; 3B05F81C53A21F2F CRC64;

Query Match 11.3%; Score 350.5; DB 16; Length 444;
 Best Local Similarity 30.2%; Pred. No. 1,6e-13;
 Matches 114; Conservative 54; Mismatches 127; Indels 83; Gaps 17;

QY 182 EXIAQXXXNDYXYITAKPEYKVMNDVARGIVKA-DVAQSSVGLYGQQLYAVADTGLDT 240
 DB 100 ESTQEMLVCKDIRKITLYNEVHALDTPAYESAQAEVITNGELTICKDVTIAYIDGT-- 157
 QY 241 GRNDSSMEAFRGKITALY-ALGRTNMANDTNGHGTAVAGSVLGNXGXTN-----KGMAPQA 295
 DB 158 -----YPHEDLGRKIKAFVDFVNOGREPYDDNGHGHGACGADANGASSDQYRGPARPA 212
 QY 296 NLVFGSINMSXGGLGSLPSNIQTLFSQAXSAGARIHTNSKGAANVAGATTTDSRNV----- 350
 DB 213 NVIGVVLNKO--GMGLSEIMOGV-----EMCIQYMEHDDPIHIISMSL 257
 QY 351 -----DDVYR-----KNDMTILFAAGNEXPNGGTSAPATKAKNAITVAGATEN 392
 DB 258 GGOALPYENEGEDPMVRYIEEAMNAGITCVAAAGNSGPAQGITASFGVSEKVIYTGALDD 317
 QY 393 LRPSFGSYADININVAOFSRSPPTKDKRIKPDVMAAGTYILSARSSLADDSF-----WA 447
 DB 318 -KDTTDREDD--VAFSSRGPTIGKPKPILAPGVIVSLRS--PNSFYDKIQKS 370
 QY 448 NHDSKAYAMGTSNATPIYAGNVQALREHFVKNRGITPSPBLKALLAGA--AD----- 500

THIS PAGE BLANK (uspto)

THIS PAGE BLANK (uspto)